FUTURE AERONAUTICAL COMMUNICATIONS

Second Edition

Edited by **Simon Plass**

FUTURE AERONAUTICAL COMMUNICATIONS

Edited by **Simon Plass**

Future Aeronautical Communications
Edited by Simon Plass

Contents

Preface

Introduction

There are well-founded concerns that current air transportation systems will not be able to cope with their expected growth. Current processes, procedures and technologies in aeronautical communications do not provide the flexibility needed to meet the growing demands. Aeronautical communications is seen as one major bottleneck stressing capacity limits in air transportation. Ongoing research projects are developing the fundamental methods, concepts and technologies for future aeronautical communications that are required to enable higher capacities in air transportation.

A study of EUROCONTROL states achievement of the aeronautical communications capacities in Europe in the next decade. Still, the main aeronautical communications is based on analog voice communication. The analog techniques are using the HF band for remote and oceanic regions and the VHF band is used for populous continental areas. There already exist aeronautical digital data links which do not increase a data rate of about 32 kbits/s but accessible satellite links. Table 1 lists available digital aeronautical data and satellite links. Moreover, the expected air traffic management (ATM) paradigm shift towards more strategic and tactical planning requires additional communication capabilities which are not provided by current air traffic control (ATC) and ATM communication systems.

	Technology	Band	Access Scheme	Modulation	Data Rate
A	VDL2 (VHF Data Link 2)	VHF	CSMA	D8PSK	31.5 kbits/s
A	ACARS	VHF	CSMA	AM MSK	2.4 kbits/s
A	HFDL (HF Data Link)	HF	TDMA	MPSK	1.8 kbits/s
S	Iridium	L	hybrid FDMA/TDMA	DE-QPSK	9.6 kbits/s
S	Globalstar	L	combination of CDMA with FDMA	Offset-QPSK	9.6 kbits/s
S	Inmarsat Swift Broadband	L	hybrid FDMA/TDMA	QPSK / 16QAM	< 450 kbits/s

Table 1. Existing digital aeronautical data links (A) and satellite links (S).

From Vision to Action

Integrating existing and future communications infrastructure in a *system of systems* is the vision of the future communications infrastructure (FCI) to enable the goals for a safe, secure and capable future ATM communications. In 2003 ICAO expressed the need of new functionalities in aeronautical communications by an evolutionary approach. The Action Plan 17 by EUROCONTROL and the US Federal Aviation Administration (FAA) developed a comprehensive view of the overall needs in 2007. From 2007 to 2009, the EU research project NEWSKY (NEtWorking the SKY) addressed these demands by launching a first feasibility study for a global airborne network design and developing initial specifications for a new aeronautical communications network based on Internet technologies (IPv6). Also the EU research project SANDRA (Seamless Aeronautical Networking through integration of Data-links, Radios and Antennas) aims at designing and implementing an integrated aeronautical communications system and validating it through a testbed and further in-flight trails on an Airbus 320. Both, the FAA and the European Commission support intensive studies in this field, namely by the NextGen and SESAR programs. Of course, additional effort is and has to be spent in the area of future aeronautical data links, i.e., satellite, L-band Digital Aeronautical Communication System (LDACS), AeroMACS to facilitate the concept of a seamless aeronautical network.

Outline of the Book

This book assembles recent research results, emerging technologies and trends in the field of aeronautical communications. The book is organized in 5 parts covering occurring trends, aspects for future aeronautical networks, the challenges for the satellite component, emerging aeronautical data links, and visions for aeronautical communications.

In the first part, *Barba & Battisti* give an insight of the recent SESAR program and the SANDRA research project including their main objectives, research activities and their collaboration. Current trends for datalink service providers (DSPs) are indentified by *Durand & Longpre* and also new emerging roles and its new players for future aircraft IT systems and associated ground components are discussed.

Future aeronautical network aspects are covered in the second part. The integration of a service-oriented architecture in an aeronautical communications environment is shown in the chapter of *Yifang et al.* whereas the system wide information management (SWIM) ground-to-ground services are extended to air-to-ground information exchange. *Muhammad & Berioli* investigate the influence of the different data traffic characteristics by ATM communications to the transmission control protocol (TCP) and detailed analysis of the ATM traffic pattern and suggestions on the system design are given. Due to safety and security requirements an aeronautical communications system is a critical infrastructure. *Pecorella et al.* give an overview of

security risks and major difference to a classical IP based network for an IPv6 based aeronautical communications network. By approaching the lower layers of a communications network, *Kissling & de Cola* study the aspects for quality of service (QoS) management in the ATM context covering also the architecture for selection of links and their interaction between technology independent and technology dependent components. The following chapter goes into more detail regarding the needed interoperability among the heterogeneous future aeronautical network. *Xu et al.* present the required functionalities for a smooth communication between the upper and lower layers of an aeronautical communications system. Finally, *Lücke & Fazli* show different design aspects for an IPv6-based aeronautical communications testbed and highlight several technical details such as IPv6 over IPv4 network transversal and robust header compression.

Already facing and upcoming challenges for the satellite component in an entire aeronautical communications system are discussed in the fourth part. An overview of the current trends, requirements and problems for a future aeronautical satellite communications link are given by *Van Wambeke & Gineste.* Reliability, low maintenance and broadband connectivity are main goals of a satellite antenna component. *Marpaung et al.* present the development of an electronically-steered phased array antenna system included optical beamforming and the resulting challenges.

At the airport, the highest density of information for flight operations exists and a secure wideband wireless communications system is proposed: AeroMACS. Also the existing VHF analog voice communications and VDL2 are a bottleneck for the ground-based aeronautical communications today. Therefore, new aeronautical data links are needed. *Fistas* describes the European view and approach on the FCI and its future proposed data links: AeroMACS; LDACS and a satellite component. An overall view on the development of AeroMACS and a first realized prototype implementation in Cleveland, Ohio, USA is given by *Budinger & Hall*. Details on the AeroMACS profile including the MAC layer design and the use of IPv6 over AeroMACS are discussed by *Ehammer et al.* The second proposed future aeronautical data link is LDACS for ground-based communications. Special focus in the following two chapters is on the LDACS1 proposal. The functional architecture and its medium access for such a link are analyzed by *Gräupl & Ehammer,* and furthermore, simulation results are provided. The closing chapter of this part handles the physical layer design of LDACS1 giving details on the frame structure, coding/modulation, out-of-band radiation and receiver design by *Gligorevic et al.*

Without any vision in research there would be no future and new goals. One vision booster is the new international platform of national aviation research organizations: International Forum for Aviation Research (IFAR). Now with 21 members, in this last part *Degenhardt et al.* introduce this new platform with a special focus on the aeronautical communications aspect. The final chapter introduces the concept of an

airborne Internet. *Medina & Hoffmann* envision the realization of an ad-hoc air-to-air Internet via the North Atlantic, giving first routing protocol strategies and simulation results.

<div align="right">

Dr. Simon Plass
Institute of Communications and Navigation
German Aerospace Center (DLR)
Germany

</div>

Part 1

Current Trends

SESAR and SANDRA: A Co-Operative Approach for Future Aeronautical Communications

Angeloluca Barba[1] and Federica Battisti[1,2]
[1]Selex Communications S.p.A.,
[2]Università degli Studi Roma TRE,
Italy

1. Introduction

The air transportation sector is currently under significant stress. The sudden decrease in demand for air based transportation after 2001 events, forced most airlines to reorganize and strength their politics by operating severe cuts and by creating strong holding to reverse the negative trend. Air traffic situation returned to pre-September 2001 levels in 2005 and nowadays the demand in aircraft operations is expected to double by 2025.

There are many concerns that current air transportation systems will be able to safely cope with this growth (FAA/EUROCONTROL, 2007; SANDRA, 2011). In fact existing systems are unable to completely process flight information in real time, and current processes and procedures do not provide the flexibility needed to meet the growing demand. New security requirements are affecting the ability to efficiently transport people and cargo. Furthermore, air transportation expansion caused community concerns on aircraft noise, air quality, and air space congestion.

3x more flights in the EU skies	10x Improved safety factor	10% environmental impact reduction per flight	50% ATM costs cut
More services, from ATS to APC, with different needs	ATM will totally rely on data communication	More bandwidth for Airlines operation & maintenance	Developing market for PAX services
New radio systems	Space and weight cost reduction	Complex network architecture evolution	Communications saturation due to traffic growth

Fig. 1. The European sky in 2025.

This scenario becomes extremely important even considering that in the past 40 years, air traffic management, indispensable for a safe flight, did not significantly progress.

A possible solution for reducing congestion problems in capacity-constrained airports has been proposed by economists through peak-load pricing. However, this solution has been rejected by both legislators and customers. At the same time, most heavily congested airports in the United States and Europe have been subject to takeoff and landing constraints, that effectively impose entry restrictions in these airports while reducing the load on air traffic control systems. The expansion of existing airports, the use of secondary airports for low-cost travels, or the creation of new huge hub increases again the awareness situation. It is therefore evident the need for a substantial change in air transportation.

In order to allow future systems to be compatible with the expected air-traffic increase, some high-level requirements on communications related aspects can be identified (Fig. 1):

- pilots' situation awareness has to be improved; this includes enhanced communication with the flight controller, monitoring communication between controllers and other aircrafts, visual look-out, and navigation (including use of maps and charts);
- airports' hosting capacity, one of the main limiting structural factors, has to be increased; there is the need to cope with the growing demand by air carriers for the use of airport facilities;
- ATS (Air Traffic Services) have to be based on reliable data communications;
- AOC (Airline Operations Control) data traffic has to increase for efficient operations;
- passengers and cabin communication systems have to be further developed in terms of robustness, reliability and re-configurability;
- safety critical information should be transmitted to the ground station in a reliable and multi-modal way; there is the need for the certainty that the information has been read, understood and implemented.
- on-board network architecture, which connects each passenger seat/crew terminal to the In-Flight Cabin server, needs convergence of protocols and interfaces.

As can be easily imagined, new technologies and operation procedures cannot be easily applied in the aviation sector: the safety, the reliability, and the effectiveness of each innovation must be deeply investigated and verified. Moreover there are standardization issues that must be taken into account when changing existing avionic equipments and procedures. Furthermore, the outcome of the security-effectiveness phase must be balanced with implementation and operational costs.

For the above mentioned reasons, both Federal Aviation Administration (FAA) in the US and European Commission in Europe, are promoting and supporting intensive studies on this field. In particular, among the initiatives supported by the European Commission, in the following the SESAR (Single European Sky ATM Research) and SANDRA (Seamless Aeronautical Networking through integration of Data links Radios and Antennas) are described in Sections 2 and 3. Since both projects are related to different aspects of the same topic, it is likely that subtasks of the projects or even some of their outcomes could overlap. We believe that from the exploitation of these synergies both projects could benefit in terms of effectiveness, costs, and overall success. In Section 4 the strategy and the undergoing efforts for revealing the projects overlaps, as well as the description of the co-operation started by the two projects on the communication aspects are reported. Finally, in Section 5 some concluding remarks are drawn.

2. SESAR - Single European Sky ATM Research

The SESAR Joint Undertaking (SJU) was created under European Community law on 27 February 2007, with EUROCONTROL and the European Community as founding members. The SESAR programme is in the framework of the Single European Sky (SES) initiative to meet future capacity and air safety needs and it is one of the most ambitious research and development projects supported by the European Community.

The mission of the SJU is to develop a modernized air traffic management system in the European air transportation sector. This system will ensure the safety and fluidity of air transport over the next thirty years (SESAR D4, 2008; SESAR D5, 2008), it will reduce the costs of air traffic management and the environmental pollution.

The key performance targets to be accomplished by 2020 (SESAR D2, 2006) are strictly related to the challenges described in the introduction:

- enable a threefold increase in capacity;
- improve safety by a factor of 10;
- reduce by 10 % the environmental impact per flight;
- cut ATM costs by 50%.

These objectives are pursued by a team of 16 members belonging to the aviation community. Furthermore, some of these members are consortiums themselves and this raises the total number of companies involved in the project to 35 units.

Due to the large spectrum of activities within SESAR, it has been partitioned in 16 Work Packages, each of them devoted to the main areas of involvement, namely (SESAR, 2011):

- Operational activities:
  - WP 4 En-Route Operations: to provide the operational concept description for the En-Route Operations and perform its validation;
  - WP 5 Terminal Operations: to manage, co-ordinate and perform all activities required to define and validate the ATM Target Concept (i.e. Concept of Operations and System Architecture) for the arrival and departure phases of flight;
  - WP 6 Airport Operations: to refine and validate the concept definition, as well as the preparation and coordination of its operational validation process;
  - WP 7 Network Operations: to cover the evolution of services in the business development and planning phases to prepare and support trajectory-based operations including airspace management, collaborative flight planning and Network Operations Plan (NOP);
- System development activities:
  - WP 9 Aircraft Systems: it covers the required evolutions of the aircraft platform, in particular to progressively introduce 4D trajectory management functions in mainline, regional and business aircraft;
  - WP 10 En-Route & Approach ATC Systems: it designs, specifies and validates the En-route & TMA ATC Systems evolutions for enhancing Trajectory Management, Separation Modes, Controller Tools, Safety Nets, Airspace Management supporting functions and tools, Queue Management and Route optimization features;
  - WP 11 Flight and Wing Operations Centres / Meteorological Services: it deals with the development of the Flight and Wing Operations Centres and with the provision and utilization of Meteorological Information services, needed to support the performance requirements of the future ATM system;

- • WP 12 Airport Systems: it encompasses all Research & Development activities to define, design, specify and validate the Airport Systems needed to support the SESAR ATM target concept;
  - • WP 13 Network Information Management System: it covers the System and Technical R&D tasks related to the Network Information Management System (NIMS), the Advanced Airspace Management System (AAMS) and the Aeronautical Information management System (AIMS);
  - • WP 15 Non-Avionic CNS System: it addresses CNS technologies development and validation also considering their compatibility with the Military and General Aviation user needs;
- • System Wide Information Management:
  - • WP 8 Information Management: it aims at developing SWIM
  - • WP 14 SWIM Technical Architecture: it is the follow-up of the SWIM SUIT FP6 Commission.
- • Transverse activities:
  - • WP 3 Validation Infrastructure Adaptation and Integration
  - • WP 16 R&D Transversal Areas: it analyzes the improvements needed to adapt the Transversal Area management system practices to SESAR as well as towards an integrated management system.
  - • WP B Target Concept and Architecture Maintenance
  - • WP C Master Plan Maintenance.

Among the described activities, one of the focal points in SESAR is the definition of the communication architecture (SESAR D3, 2007). The SESAR ATM concept requires advanced data communication services and architectures able to support specific features such as: 4D-trajectory management in order to be able to update and revise the Business Trajectory of the aircraft, ASAS separation to allow the crew to perform some tasks related to separation or spacing, thereby reducing the workload of the controller, and SWIM operations, as described in the following (SWIM, 2011).

The SWIM (System Wide Information Management) is one of the focal aspects in the approach proposed in SESAR. It aims at the replacement of data level interoperability and closely coupled interfaces with an open, flexible, modular and secure data architecture that is able to support users and their applications in a transparent and efficient manner, see Fig. 2.

(a)   (b)
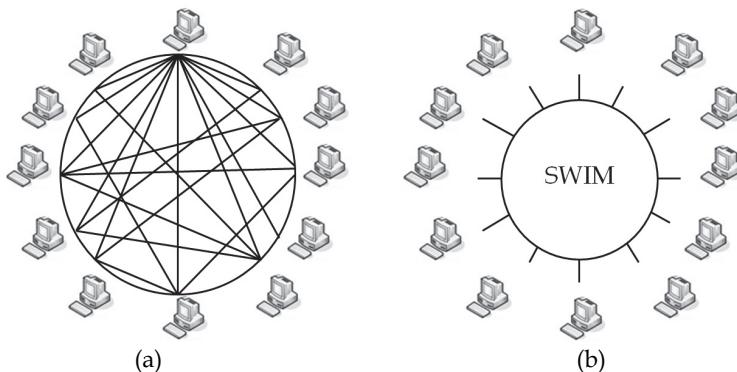
Fig. 2. Actual architecture (a) and architecture based on the SWIM concept (b).

SWIM will be used for enabling data sharing between ATM services across the whole European ATM system. Even if a complete definition of the SWIM has not yet been reached, interoperability and standardization are key elements and SWIM is expected to play an important role in the revolution of the aeronautical scenarios.

Toward this direction, ATM stakeholders will cooperate in the development of SWIM requirements, prototypes, roadmaps, and deployment plans.

In particular SWIM will provide benefits to:

- pilots during takeoff, navigation and landing operations by guaranteeing a reliable communication with the air traffic controller who will give support and instruction based on data collected and validated from different sources;
- Airport Operations Centers, managing departures, surface movements, gates and arrivals, building schedules, planning flight routings and fuel uplift, ensuring passenger connections and minimizing the impact of delays;
- Air Navigation Service Providers (ANSPs), organizing and managing the airspace over a country and with Air Traffic Services – managing air traffic passing through their airspace;
- Meteorology Service Providers for weather reports and forecasts;
- Military Operations Centers, planning missions, securing airspace during training operations, fulfilling national security tasks.
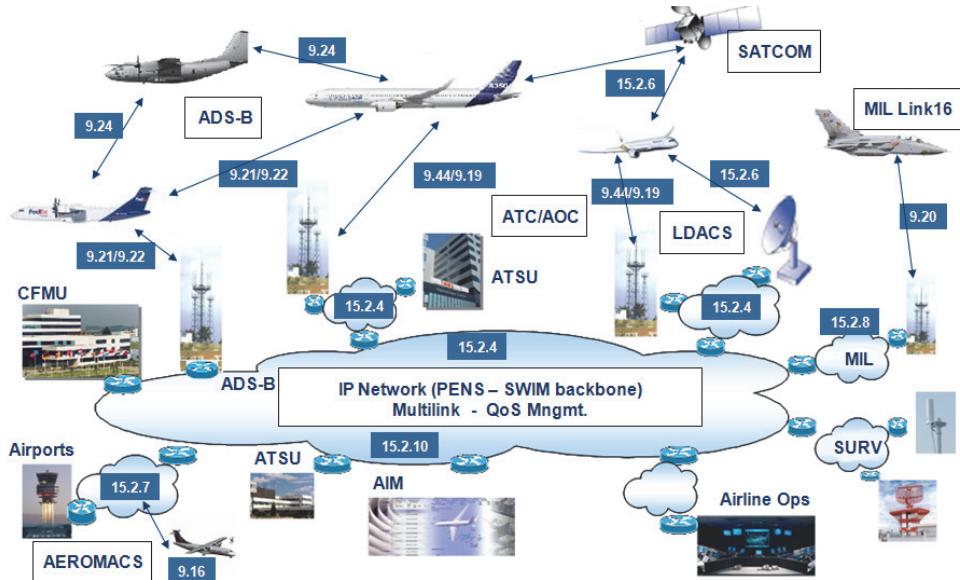
Fig. 3. Communication architecture.

A reliable and efficient communication infrastructure will have to serve all airspace users providing the appropriate Quality of Service (QoS).

It is useful to underline that QoS is a complex concept but in an air/ground communication link it can be roughly measured by using three parameters: communication delay, data integrity, and system availability. As demonstrated by many studies, the real quality

requested by a communication system is strictly dependent on the particular service and operational scenario. The relative importance of the identified parameters is determined according to the particular application: for instance the real time communication between the pilots and air traffic management system in high density traffic area requires the delay to be as short as possible, to have high data integrity, and high service availability. At the same time, the data link adopted for delivering or predicting meteorological conditions for low-density airspace, might accept longer delays, less integrity, and lower availability. In a modern communication scenario, other parameters can influence and contribute to the overall QoS, as the fulfillment of the authentication, authorization, and accounting requirements, the customer satisfaction, and so on.

Last, it should be also mentioned that the provisioning of QoS strongly reflects on service costs: the exact estimate of the QoS required by the application may avoid increased-unjustified costs thus preventing the service from being used.

As shown in Fig. 3, the overall QoS will be guaranteed to the particular application, through a communication scenario involving both mobile and fixed entities. While the definition and the provisioning of QoS in fixed communication systems has been studied and achieved during the last fifty years, the same goal for mobile communication is still far to be reached. The noisy nature of the communication media itself, together with security concerns, and the need of fusing different communication approaches, can be considered a big challenge for present and future communications.

In SESAR, the mobile part of this infrastructure will be based on a multilink approach, composed of different sub networks:

- a ground-based L-band line of sight data link as the main system in continental airspace;
- a satellite-based system (in cooperation with the European Space Agency) to serve oceanic airspace whilst complementing ground-based data link;
- a system dedicated to airport operations, derived from WiMAX, providing a broadband capacity to support the exchange of a significant amount of information;
- to allow interoperability with military operations, a gateway is being defined to interconnect the ATM system and the military Link 16 system.

At architectural level this future infrastructure will incorporate the legacy networks as VHF/VDL and the growth capability toward eventual future evolutions (SESAR D3, 2007).

3. SANDRA - Seamless Aeronautical Networking through integration of Data links Radios and Antennas

SANDRA is a project partially funded by the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement n. 233679 (SANDRA, 2009; SANDRA web, 2011).

This project aims at the definition, the integration, and the validation of a reference communication architecture, SANDRA Airborne Communication Architecture, directly related to the Service Oriented Avionics Architecture envisaged in the Future Communications Study (FAA/EUROCONTROL, 2007).

The SANDRA consortium consists of 30 partners from 13 countries across Europe composed by industrial partners, research organizations, universities and highly specialized SME (Small and Medium Enterprises).

This project defines specific targets to be accomplished in order to face the new needs in the aeronautical transportation sector. These can be achieved by integrating existing and novel heterogeneous communication media into an overall architecture able to:

- provide and manage seamless service coverage across any airspace domain and aircraft class;
- support the increasing trend of the service market and to enable easy plug-in of future radio technologies by means of a modular and reconfigurable approach;
- be upgraded, easily reconfigured and independent on the specific radio technology used;
- be distributed on ground-base and airborne sub-networks thus ensuring full interoperability.

The novelty of the SANDRA approach consists in pursuing integration at different levels:

- service integration: integration of a full range of applications (ATS, AOC/AAC, APC);
- network integration: based on interworking of different radio access technologies through a common IP-based aeronautical network whilst maintaining support for existing network technologies (ACARS, ATN/OSI, ATN/IPS, IPv4, IPv6);
- radio integration: integration of radio technologies in an Integrated Modular Radio platform;
- antenna integration: L-band and Ku-band link antennas will be used to enable an asymmetric broadband link;
- WiMAX adaptation for integrated multi-domain airport connectivity.

Ultimately, SANDRA pursues the architectural integration of aeronautical communication systems using well-proven industry standards like IP, IEEE 802.16 (WiMAX), DVB-S2, Inmarsat SwiftBroadBand, a set of common interfaces, and standard network protocols having IPv6 as final unification point to enable a cost-efficient global and reliable provision of distributed services across all airspace domains and to all aircraft classes.

The SANDRA validation activity will show the ability of the proposed integrated architecture to easily reconfigure and adapt for the flexible implementation of new communication services.

In terms of overall working structure, SANDRA is structured in eight Sub Projects (SP), each one dedicated to a specific aspect (Fig. 4). In more detail, SP1 is related to the project management. In SP2 the "top-down" approach, the scenarios, the overall framework and architecture are defined and developed. SP2 is therefore the central conceptual integration activity in the project.

Moreover, the project structure clearly reflects the focus on the four major SANDRA elements, namely:

- Seamless Networking (SP3): SANDRA networking solutions are designed to allow integration and interoperability at different levels, with IPv6 as final unification point (target 2025 and beyond):
    - Link level: Interworking of different link technologies (ground-based, satellite-based, airport systems as main streamline for validation, air-to-air MANET as long term extension);
    - Network level: Interoperability of network and transport technologies (ACARS, ATN/OSI, IPv4/IPv6) to ensure a realistic transition;
    - Service level: Integration of operational domains (ATS, AOC/AAC, APC). This integrated networking approach of SANDRA is a key enabler for the efficient implementation of a range of applications addressing the ACARE SRA2 objectives,

that are the development of a customer oriented, time efficient, cost efficient, green and secure air transport system. The SANDRA networking solution is also a key enabler for most SESAR Key Performance Areas (KPAs) as defined in SESAR deliverable D2 (SESAR D2, 2006).

- Integrated Modular Radio (SP4): As all of the radios will never be used simultaneously, there is opportunity to build a new radio system in which each single radio element can be independently reconfigured to operate in a specific radio link as required. This will considerably reduce the amount of radio sets in the vehicle thus reducing weight and consequently the operational costs. Furthermore, the different type of radios will be replaced by one software-reconfigurable equipment, thus simplifying spares and maintenance operations. The adoption of software defined radio based equipments greatly simplifies the seamless transition from one link to another that is the goal of the networking aspects of the SANDRA programme. Another progress will be the ease of integrating the radios into the overall avionics system as they will have identical interfaces both in software and hardware terms. Also, pilots' workload will be reduced, not only because the radio operation will be simpler, but also because the seamless networking capability of the overall system removes the need for the pilots to manually select radios. Finally, just as IMA (Integrated Modular Avionics) allows the avionics to be updated with new applications by means of software change only, similarly the IMR will allow future communications waveforms and protocols to be updated by software alone.

- Integrated Antennas (SP5): A key requirement for future aeronautical communications systems is the provision of broadband connectivity within aircraft cabins at an affordable price. One of the key enablers is an electronically steered Ku-band phased array. Ku-band phased arrays in which the same elements are used for both transmission and reception are not possible with mature technologies. Consequently, for Ku-band two arrays would be required. Given the undesirability of increasing the number of antennas, such a solution is not acceptable in the market place. However the amount of data agglomerated over a range of passenger services (VoIP, Web, Email, SMS, MMS) and over a range of flights (SH, LH), is highly asymmetrical, with the inbound traffic being about 5 times higher than the outbound. The inbound traffic requires the availability of a broadband Ku-band antenna in receive mode only, which is feasible. A further benefit of a receive only system is that the beam width restriction to avoid inadvertent irradiation of other satellites can be reduced; this is a particularly useful amelioration as it means that the phased array can be used (maybe, at slightly reduced data rates) at low elevation angles, where the beam tends to flatten out. The other key element in making this link work in totality is the asymmetrical networking aspect of using different bearers for the forward and return link. A dedicated signaling system will be developed in SANDRA as a general concept of IP based data exchange using asymmetric bearers.

- Airport Connectivity with WiMAX (SP6): Airports are the nexus of many of the Air Transport transformation elements to achieve air traffic management (ATM), security, and environmental goals. Accordingly, the sustainability and advancement of the airport system is critical to the growth of the air transportation system. To enable these progress and vital concepts like "fast turnaround", SANDRA will define and validate a new generation Airport Wide Area Network, supporting a large variety of vital

aeronautical applications and services. This network is envisioned as a high-integrity, safety-rated, wireless network, with mobile terminals on the ground, on aircraft as well as other surface vehicles and improved information distribution.To respect the high level of security needed in the aeronautical environment, SANDRA will consider security measures into the design at all layers. The new system will be based on WiMAX standards for ATS/AOC communications and will provide lower cost, safer and more efficient airport surface operations, compliant with SESAR/FCI recommendations, giving strong input in know-how to the SESAR initiative.

Finally, the overall SANDRA system encompassing prototypes of the integrated router, the integrated modular radio and the integrated antenna will be validated in SP7 in laboratory test-bed and in-flight trials using the ATRA (Advanced Technology Research Aircraft) an Airbus A-320 of the German Aerospace Center (DLR)..

Exploitation and dissemination activities are coordinated in SP8 where a particular emphasis is put on the coordination with SESAR.
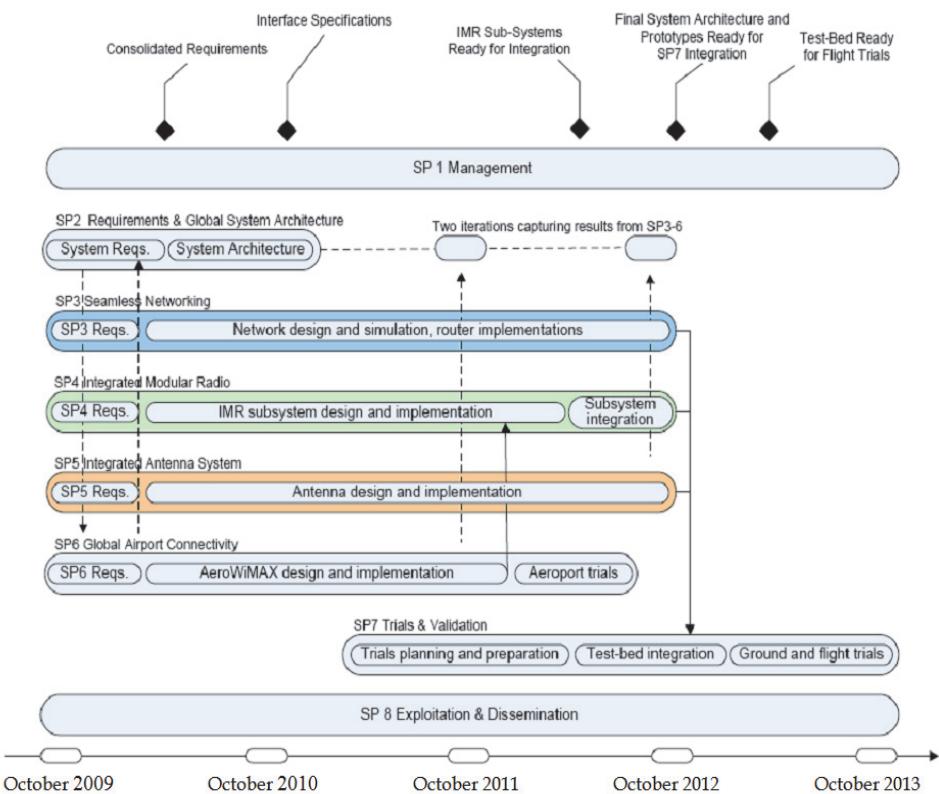
Fig. 4. SANDRA study logic and key milestones.

SANDRA shall use and build upon other project results at different levels. In some cases only the knowhow about emerging/future communication technologies, requirements and interfaces will be reused, i.e. no hardware and software will be carried over. In other cases,

reuse will be done also of platforms and development items. This interaction is summarized in the following list:

- Future Communication study for technologies requirements (voice and data), description and conclusions on next steps (such as AeroWiMAX, need for new Satcom, L-DACS 1/ 2);
- ESA IRIS study for new satellite ATM data link;
- Studies and reports from SESAR, COCR and NASA MCNA for requirements, concept and architectures;
- Previous EC projects such as Wireless cabin, E-CAB, MOWGLY, ANASTASIA, NEWSKY for requirements and architectures (system architecture and on-board architecture) for ATM and APC communications;
- ICAO WG-I (Internetworking) and corresponding ATN/IPS documents for IP protocols for safety related communications, IETF for possible evolution of IP protocols for aeronautical applications;
- EUROCONTROL IP study and NASA MCNA studies for architecture and protocols;
- SWIMSUIT EC project and SESAR for SWIM and SOA concepts for ATM networks, and their impact on the architecture;
- NEWSKY for preliminary integrated network architecture design supporting ATS, AOC and APC;
- SESAR Master Plan and NEWSKY for Road Map and transition aspects;
- ANASTASIA the concept of a hybrid dual-frequency L/Ku antenna, preliminary dual band patch designs and first prototype radio architecture with a completely Software modem;
- Future Communication Studies and E-CAB for AeroMACS requirements (safety critical and not) ;
- B-AMC for consideration of load sharing between terrestrial and satellite communication systems; performance enhancement of terrestrial system;
- NATALIA ESA project and Flysmart (NL) for design and technologies related to full electronically steerable Ku-band antenna with broadband optical beam forming;
- SAFEE for risk assessments, security analysis for the communications domains, Security for existing architectures;
- Requirements from Logistic & Health Management Oriented projects to assure SOA interoperability (i.e. ECAB).

Fig. 5 shows the existing connections between SANDRA and the above-mentioned projects and standardization activities.

4. The SANDRA-SESAR collaboration

4.1 Context

In 2009 the SANDRA Consortium, the DG Research, and the SESAR Joint Undertaking established a co-operation. The aim of this agreement is to share resources with a flexible approach and to provide the European community with benefits obtained both by individual and integrated outcomes of the two projects. Moreover, it is fundamental for Europe to show a unified approach in the aeronautical field.
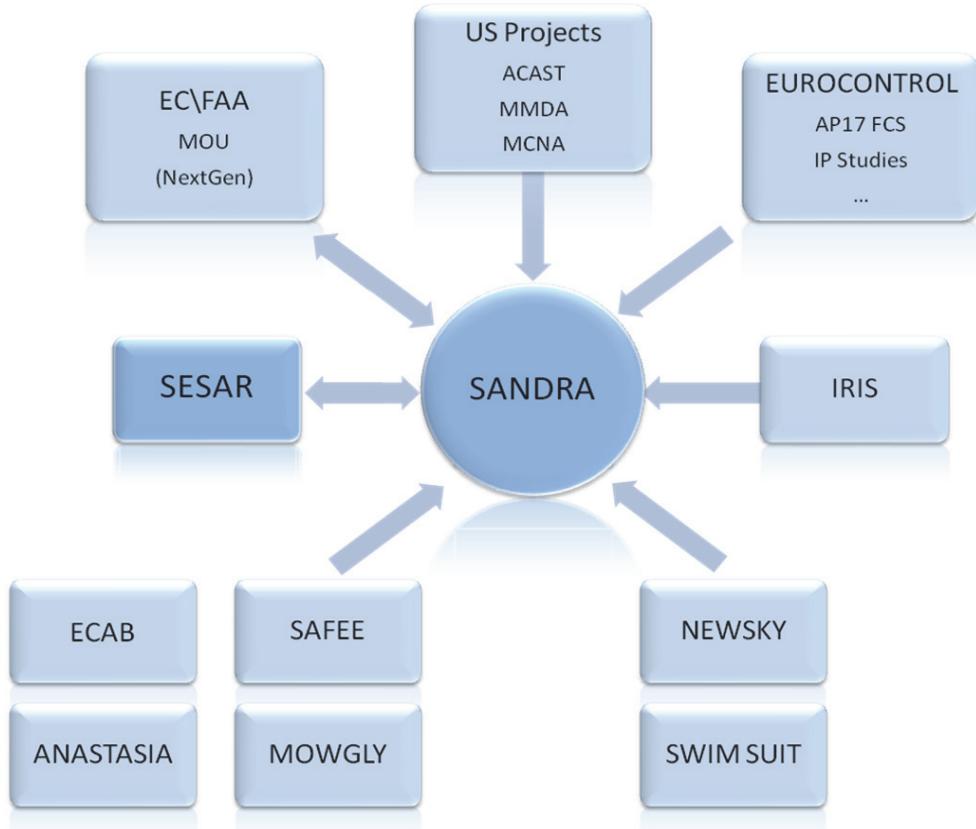
Fig. 5. Relationship between SANDRA and other projects and activities.

To achieve this ambitious collaboration, a set of Work Areas (SANDRA, 2011; SESAR D6, 2008) were identified:
- definition of requirements,
- multilink and QoS management,
- flexible communication avionics,
- airport wireless communication systems,
- architecture, networking, and SWIM airborne.

The proposed approach reflects the need to optimize the common efforts. This is achieved by gradually exploiting the results obtained by the single research programmes also considering their peculiarities as time scheduling, final objectives, and required competencies.

Fig. 6 shows the tight connection between projects and studies in the SANDRA-SESAR co-operation that will be analyzed in the following sections.

Similarly, in USA the Federal Aviation Authority has proposed the NextGen project. The goal of this project is to fuse different competencies in the field of National Airspace System and projects for realizing a more convenient and dependable travel system, while ensuring the safety and security of the flight.

According to the project developers, the outcome of this cooperation will optimize the economic aspects, the impact on environment (pollution), the information delivering and exploitation, the safety management and prevention, the interaction among the different actors (users, travel companies, airports, cargo systems, ground transportation and services), and will increase the overall security.
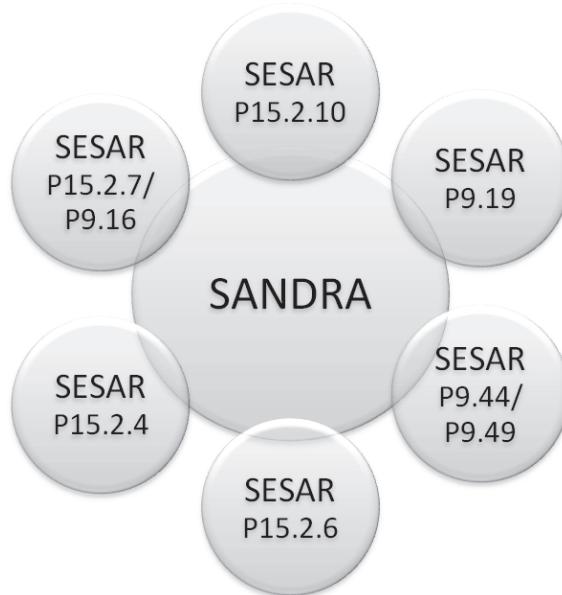
Fig. 6. List of feeder projects, studies and initiatives.

4.2 Overall concept and architecture comparison

In order to understand the relation between the programmes and their possible synergies, the conceptual differences in the approaches has been investigated. Several outputs of the SESAR Definition Phase (2006-2008) were used as inputs for the requirement definition and functional architecture design. In particular:

- Deliverable 3 - 'Future ATM Target Concept' (SESAR D3, 2007) describes the main concept of operations, the architecture for future ATM System, the set of identified enabling technologies, the outline of total costs, and the positive outcomes of the feasibility study;
- Deliverable 4 - 'Deployment Sequence - Develop Options and Select 'Best' Practices' (SESAR D4, 2008) contains the confirmation of feasibility (technical, financial, institutional, etc.), the development of options and the recommended approach for the deployment phase, and the definition of deployment packages (transition from legacy systems/framework);
- Deliverable 5 - 'ATM Master Plan' (SESAR D5, 2008) details the plan of actions that all organizations need to implement, the possible outcomes to be used in future business plans, RT/D plans, risk assessment studies, and it envisages future management processes.

The SANDRA system interfaces have been defined taking into account on the Air-to-Ground interoperability requirements specified in SESAR. The relation between SANDRA and SESAR is extremely important since SANDRA aims at defining an architecture that is compliant with SESAR IP3 communication baseline as exposed in SESAR WP2.5/D4 'Technology Assessment' (SESAR D4, 2008)

For what concerns the technological aspects, a detailed analysis has been conducted to confirm SANDRA's fundamental coherence with the SESAR concept.

Following a detailed analysis of the two projects, significant correspondences have been identified in five macro areas concerning Software Defined Radio (SDR) Architectures, Integration, Network architecture, Security, and Airport Wireless LAN.

Those aspects are highlighted in Table 1- Table 5.

Table 1 reports the approach followed by the two projects on the SDR Architectures topic. For example it can be noticed that in both projects the flexibility in radio resources exploitation is a key investigation element. To achieve the desired flexibility both projects envisage the use of SDR.

SDR Architectures	
SESAR	SANDRA
Software defined radios are available for avionic integration and global interoperability.	Minimization of the radio hardware equipment by reconfigurable avionic radios. Flexible radio resources: key enabler in the planning of the new links being undertaken by SESAR.
Flexible development and rapid evolutions (e.g. through SDR technology) are desirable	A scalable architecture that allows a flexibility in the radio resources to be added to the aircraft according to the number of users, availability and integrity requirements.
SESAR is mainly focused on AOC and ATC operations.	The main objective of SANDRA is the flexible integration of networks and technologies envisaging the convergence of ATM, AOC, APC communications for radio and routing in any operational phase.
Additional data link performance is required to support advanced services such as 4D trajectory management and increasing traffic growth. A dual link system is likely to be needed.	The Integrated Modular Radio reconfigurability is a key factor enabling efficient implement the dual link concept. SANDRA will define and implement a network layer and the various data link layers to guarantee independence of routing from links, support of critical functions over low-bandwidth links and link topology, availability, quality will be indicated to the router.

Table 1. Relationship on SDR Architectures.

Table 2 is related to the integration concerning the management of flexible aeronautical routing. Also in this case both projects are concerned with radio exploitation for an effective and reliable routing path delivery.

Integration	
SESAR	SANDRA
Integration of both continental and oceanic routing with radio capabilities.	The main objective of SANDRA is the integration of networks and technologies envisaging the convergence of ATM, AOC, APC Communications for radio and routing in any operational phase.

Table 2. Relationship on integration.

Network architecture	
SESAR	SANDRA
The transport and internetworking layers will have to be meet QoS requirements and safety and performances needed by ATS.	SANDRA enhanced routing protocols will manage all aircraft mobility and prioritize traffic end-to-end in compliance with QoS requirements. Policy based routing will be available to enable the selection of the appropriate link for every data flow. Better integrity and safety-of-flight due to the reuse of all available connections in critical conditions. SANDRA Network management will operate and integrate all the communications technologies.
Sharing with other uses (such as AOC) is envisaged.	SANDRA envisages the architectural convergence of communications domains and is fully in line with and for some aspects exceeds the SESAR vision.
Could be based on improvements to ATN or a specific augmented IP layer.	The SANDRA IPv6 orientation and the development of interoperability concepts are fully in line with the SESAR vision. Interfacing ATN networks will be considered in specific activities.

Table 3. Relationship on information network architecture.

Table 3 shows the impact of QoS and security requirements on the Network Architecture. This fundamental task is approached by both projects by designing a IPv6-based communication system allowing the interoperability among different domains.

Table 4 analyzes the approach carried out on the security aspect. The presence of a security system architecture based on encryption and AAA (Authentication, Authorization and Accounting) services, is investigated in both projects.

The correspondences in the airport wireless LAN for airport usage are detailed in Table 5. In both architectures, a tuning of the communication standard 802.16 (IEEE 802.16, 2009) is used for optimizing the communication link.

Security	
SESAR	SANDRA
Security Applications like firewalls, encryption, and authentication will be needed.	SANDRA will address an information security (INFOSEC) architecture to guarantee the separation between the different domains on the SANDRA system architecture.
Resistance to voluntary interference is analyzed.	SANDRA will consider link encryption, access authentication, accounting and link protection at RF level (anti jamming frequency hopping, etc…).

Table 4. Relationship on secure data exchange.

Airport Wireless LAN	
SESAR	SANDRA
Terrestrial data link for airport surface supporting ATS and AOC with QoS management. Initial 802.16 for AOC may provide a learning platform to define the suitable ATS surface datalink operating in a protected band.	SANDRA will define the optimum WiMAX profile, based on multiple representative airport surface propagation characteristics. The maximization of spectral efficiency, cell-planning, the management of interferences and the minimization of airport base stations, the study of infrastructure and on-board WiMAX complexity and cost, will be addressed. Traffic flow monitoring will enable fine-tuning of the WiMAX profile to optimize the waveform to all airport propagation characteristics.

Table 5. Relationship on terrestrial point to point data link for airport usage.

Finally, as shown in Table 6, there is a strong correlation between the expected SANDRA outcomes (SP3 to SP7) and the communications enablers identified in SESAR D4 for implementation packages (IPs) 2 and 3.

The most correlated topic is the New Airport Datalink. It involves with major impact the SESAR IP2 with SANDRA SP3, SP4, SP6, and SP7. Even if the connection impact is not as strong as in the above mentioned cases, SANDRA Sub –Projects are related to SESAR IP2 and IP3 also on the Enhanced VHF Digital Mode 2 (VDL2) Air/Ground Data Link investigation, the Ground IP Network, the Digital Air-Ground Voice, and the Air to Air Datalink.

From the above considerations it is evident that the exploitation of redundancy between the two projects can result in optimization of both efforts and outcomes.

Despite the mentioned interactions, SANDRA and SESAR present a different approach to the architecture: SANDRA proposes an integration of information domains characterized by

safety needs, and it aims at maximizing the reconfigurability and minimizing the costs of avionic platforms. On the other hand SESAR is more oriented to the ATM field.

		SANDRA SP3	SANDRA SP4	SANDRA SP5	SANDRA SP6	SANDRA SP7
SESAR IP2	Enhanced VHF Digital Mode 2 (VDL2) Air/Ground Data Link	X	X	-	-	X
	New Airport Datalink	X	X	-	X	X
	VoIP for Ground Segment of Air-Ground Voice	-	-	-	-	-
	Ground IP Network	O	-	-	-	O
	High performance Air Ground Datalink	X	X	-	-	O
SESAR IP3	Digital Air-Ground Voice	O	-	-	-	O
	Air to Air Datalink	X	-	-	-	-

Table 6. SANDRA expected impact on SESAR IP2 and IP3 communications enablers. X stands for 'major impact' and O for 'impact'.

Based on the analysis of these different points of view, it has been agreed that SANDRA will contribute to SESAR Development Phase providing its technological outcomes and preliminary work.
SANDRA will also define the standardization activities for ATM and the exploitation efforts that will be finalized by SESAR.
This synergy is possible because SANDRA architectural integration concept of different domains is fully compatible with SESAR. As mentioned before it maximizes the reconfigurability aspects and it minimizes the costs of avionic platforms thus representing a possible evolution for the SESAR system.

4.3 Working approach

In the previous sections the similarities between the two projects have been highlighted. As a consequence, in order to merge SANDRA and SESAR work plans, several collaboration working areas have been identified (Section 4.4).
The adopted procedure for the integrated working approach is based on the following guidelines:

- for each working area, an agreement on a common work plan is established and used by both teams at working level. This is crucial for synchronization; Fig.7 shows the foreseen interaction timeline between the projects;
- on a regular basis (e.g. every six months) meetings are scheduled for assessing progress, reviewing common work plans, analyzing eventual variation on scopes or contractual agreements such as SANDRA Description of Work and SESAR Project Initiation Reports.
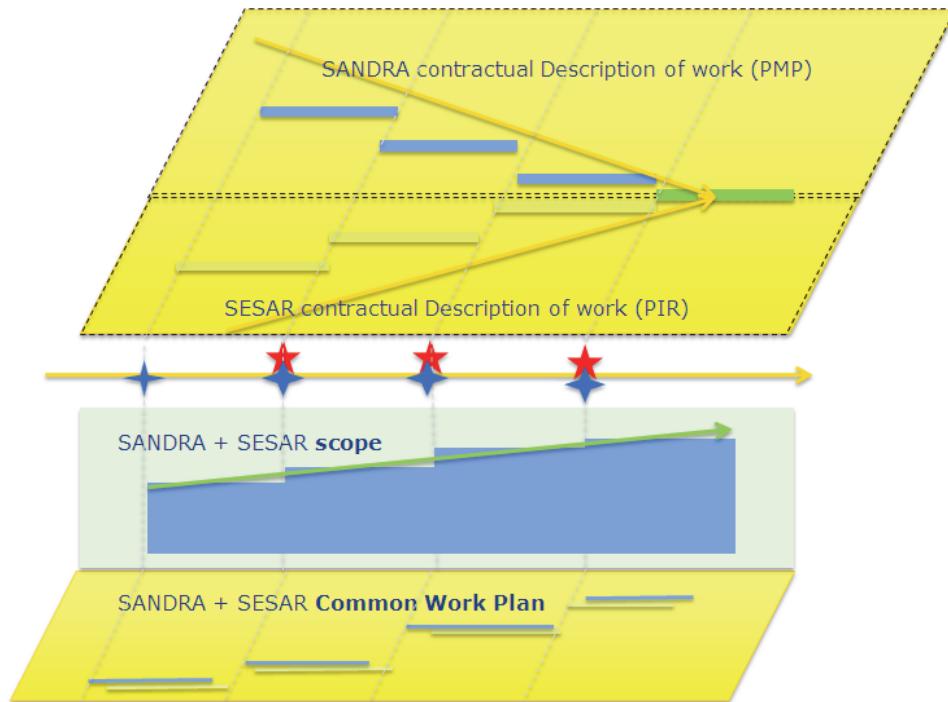
Fig. 7. Timeline of the interaction between SANDRA and SESAR.

Concerning this agreement, the European Community board showed its support to the co-operation between the projects but it required the fulfillment of the final goals of each single project: SANDRA and SESAR can exploit the beneficial aspects of sharing selected tasks but this interaction does not have to interfere with the finalization of the objective of each individual programme.

Moreover the definition of such agreement lead the two involved projects to foresee the possibility of project modifications through a Change Request Process.

The operative approach for work sharing depends on the particular working areas:

- activities can be shared between SANDRA and SESAR teams (e.g. airport communication system),
- results can be shared (input-output mode) when activities are time-sequential,
- a mixed approach can be adopted: input-output mode at the beginning and activity sharing during the following phases.

It has been agreed that the approach to be used will be identified on a case-per-case basis depending on the particular conditions.

It is also important to notice that for each working area, the common work plan has to address at least the following items:

- Work Breakdown Structure (WBS): to efficiently synchronize the common work packages and the technical activities that have to be carried out;
- Organizational Breakdown Structure (OBS): needed to share and organize the responsibilities for project management;
- Information workflow: it is necessary for the correct co-operation execution. It is mainly based on documents exchange but also on dedicated meetings;
- Respect of Intellectual Property Rights (IPR): in order to ensuring the non infringement of SANDRA and SESAR IPR rules and by analyzing case-per-case the presence of potential issues regarding the intellectual IPR violation;
- Non Disclosure Agreement (NDA): the involved parties agree to protect the confidentiality of the information disclosed in the common work.

4.4 Areas of collaboration

Starting from the analysis performed in Section 4.2 in which the architecture comparison is performed, nine common working areas have been identified and listed in Table 7. The corresponding SANDRA SPs and SESAR Projects are highlighted.

Working Area	SANDRA SP	SESAR Project
Requirements	SP2	15.2.4
Multilink management	SP3	15.2.4
Networking and architecture	SP3	15.2.4
Airport WiMAX comms	SP6, SP7	9.16, 15.2.7
QoS management	SP3	15.2.4
Software Defined Radio	SP4	9.44
Trials	SP7	All
Standardisation	SP8	All
Service Integration	SP2, SP3	9.19, 14
Airborne Infrastructure	SP2	9.44, (9.49), 9.19

Table 7. Identified working areas.

4.5 Cooperation with U.S.

Europe and the United States, being the main actors in the airspace field, are developing modernized ATM systems and their interoperability is of primary importance. However, as previously mentioned, the European aeronautical scenario is not unified and therefore there is the need for a common view.

The existence of a unified approach in the European countries, ease the relationship with the International Civil Aviation Organisation's (ICAO) Global ATM Operational Concept (ICAO, 2011). This connection is of primary importance because ICAO provides governments and industry with objectives for the design and implementation of ATM and it supports communication, navigation and surveillance systems.

To this aim a strong effort has been devoted in the SANDRA/SESAR collaboration framework in order to share the technology and procedures under development with ICAO and aviation authorities, as well as standardization bodies such as EUROCAE (EUROCAE, 2011) and RTCA (RTCA, 2011). A practical example is the coordinated effort in exchanging information with the relevant U.S. Stakeholders on the airport wireless technologies. Currently the definition of a common standard is foreseen and SANDRA and SESAR participants actively co-operate in this investigations.

4.6 Open issues

Some open issues remain, in particular when dealing with the relationships between two programmes that present different objectives, timescales and extension:

- definition of rules for solving possible project conflicts,
- definition of sharing information methodology,
- definition of a co-operating team,
- selection of an executive board.

These issues are still open and a final solution has to be found. In the next future the co-operation will lead to the definition of rules in order to maximize the synergy and the impact of the programmes on the global research and on the development in the field of aeronautical communications.

4.7 Case study: airport wireless communications

During a preliminary analysis it resulted that the operating Airport communication systems was effective and that it could be used as a pilot for this coordinated approach. The main goal of this working area is the definition and implementation of an IEEE 802.16e (IEEE 802.16e, 2009) dedicated wireless network profile, specifically tailored to aeronautical airport applications. This system is named AeroMACS and it is envisioned to operate in the 5091-5150 MHz band assigned by WRC 2007. As can be easily understood, the development and standardization of a unique profile for both European Union and United States is strongly desirable.

During the analysis, the following objectives for the common work were identified: requirements definition (including security aspects), profile definition, channel modeling, tools specification, standardization processes and trials set up.

In this process a team composed by representatives from a number of relevant sub projects was identified:

- SANDRA:
  - SP6: its main objective is the design of an aeronautical standard based on IEEE 802.16e (WiMAX) and that will use the MLS sub-band for airport surface operations, following the Future Communications Study technology assessment recommendations.
  - SP7: a test-bed for validation purpose of the overall SANDRA concept and architecture will be implemented in this SP. On-ground and in-flight trials will be used to show and prove the integrated SANDRA approach and its benefits with respect to existing aeronautical communications systems based on single radio technologies, thus incapable to overcome limitations of individual radio access systems, e.g. limited coverage of direct A/G data links, high delay of satellite systems, etc.

- SP8: this SP is devoted to the investigation of key themes from FCS to speed up standardization and adoption processes, to develop transition and exploitation concepts integrated with SESAR approach and to contribute technological results and preparatory work envisaging standardization and exploitation effort being finalized in SESAR.
- SESAR:
  - 9.16 - New Communication Technology at Airport: this is designed to define, validate and demonstrate a technical profile and an architecture for a new generation of airport surface system to enable advanced surface CNS systems and improved information distribution and provide lower cost, safer and more efficient airport surface operations
  - 15.2.7 - Airport surface Data Link: its main objective is to define, validate and demonstrate a new surface communication link that will be based on the IEEE 802.16e standard, adapted for ATS/AOC communications and compliant with FCI recommendations.

The team work activities are focused on the definition of virtual work packages that specify the activities to be completed and a planning to avoid eventual overlapping. In addition the process has been identified:

- the 'prime' of each activity: that is the responsible for all technical and management issues related to that activity;
- the role of each participant: in order to optimize the common efforts,
- a list of potential risks (e.g. timeframe) that can be found in the work development and consequent recovery actions.

5. Conclusions

In 2009 the SANDRA Consortium, the DG Research and the SESAR Joint Undertaking established a collaboration for sharing resources and for providing the European community with an extensive set of results. Since both projects are related to different aspects of the same topic, subtasks of common interest have been identified.

In particular five Work Areas were highlighted: requirements, multilink and QoS management, flexible communication avionics, airport systems, and architecture, networking, SWIM airborne.

In this chapter a detailed description of the two projects and their co-operation was presented (together with a case study) to show and highlight interactions between programmes, the working approach, and the co-operation with USA.

In this chapter it has been shown that research and industrial programmes can efficiently collaborate and that the key objective expected is the coordination of the effort in a sector, the Aeronautical Communications, which presents an enormous competition among few well harmonized stakeholders. This is an important result for resource optimization reasons, and for investigating a novel way to maximize the impact of the advanced research in the European environment.

6. Acknowledgment

233679. The SANDRA project is a Large Scale Integrating Project for the FP7 Topic AAT.2008.4.4.2 (Integrated approach to network centric aircraft communications for global aircraft operations). The project has 30 partners and started on 1st October 2009.

7. References

EUROCAE, The EURopean Organisation for Civil Aviation Equipment (2011). Information available from http://www.eurocae.net/

FAA/EUROCONTROL, Cooperative Research and Development Action Plan 17 - Future Communication Study, Final Conclusions and Recommendations Report (2007). Version 1.1. Available from http://www.eurocontrol.int/communications/gallery/content/public/documents/AP17_Final_Report_v11.pdf

ICAO, International Civil Aviation Organization (2011). Available from http://www.icao.int/

IEEE 802.16, Broadband Wireless Metropolitan Area Networks (2009). Available from http://standards.ieee.org/about/get/802/802.16.html

IEEE 802.16e, IEEE 802.16e Task Group (Mobile WirelessMAN®) (2006). Available from http://www.ieee802.org/16/tge/

NextGen, Next Generation Air Transportation System (2011). Information available from http://www.faa.gov/nextgen/

RTCA, RTCA, Inc. (2001). Information available from http://www.rtca.org/

SANDRA, Seamless Aeronautical Networking through integration of Data links, Radios, and Antennas - Grant Agreement n. 233679 (2009).

SANDRA web, Seamless Aeronautical Networking through integration of Data links Radios and Antennas (2011). Information available from http://www.sandra.aero/

SESAR D1, Air Transport framework: The current Situation, Version 3.0 (2006). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0602-001-03-00.pdf

SESAR D2, The performance Targets, DLM-0607-001-02-00a (2006). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0607-001-02-00a.pdf

SESAR D3, The ATM Target Concept, DLM-0612-001-02-00a (2007). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0612-001-02-00.pdf

SESAR D4, The ATM Deployment Sequence, DLM-0706-001-02-00 (2008). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0706-001-02-00.pdf

SESAR D5, The SESAR Master Plan, DLM-0710-001-02-00   (2008). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0710-001-02-00-D5.pdf

SESAR D6, Work Programme for 2008-2013, DLM-0710-002-02-00 (2008). Available from http://www.eurocontrol.int/sesar/gallery/content/public/docs/DLM-0710-002-02-00-D6.pdf

SESAR, Single European Sky ATM Research (2011). Information available from
        http://www.sesarju.eu/
SWIM,   System Wide Information Management (2011). Information available from
        www.swim.gov

Handling Transition from Legacy Aircraft Communication Services to New Ones – A Communication Service Provider's View

Frederic Durand and Luc Longpre
Société Internationale des Télécommunications Aéronautiques (SITA)
France and Canada

1. Introduction

The internet hasn't just changed the way we communicate – it has irrevocably altered the way we live, work, consume and spend our free time. Now, technologies that use the internet protocol (IP) in the air transport industry will lead to a similarly dramatic transformation – this time to aircraft operations, whether on the ground or in-flight. Nowhere will this transformation be more evident than the way aircraft communicate. Passengers are already benefitting from this revolution: Passenger connectivity systems already provide internet access, cellular telephony communications while aircraft is in flight.

Today, most options for onboard and external data communications offer limited capacity and versatility. This explains why the amount of information that can be exchanged is restricted to short messages, mainly in predetermined formats. It also explains why a proportion of communications are still conducted over voice. What's more, while the plane is on the ground, there are currently only limited ways of cost-effectively transferring large amounts of information and many of these involve manual downloads and physical storage media. Such constraints have had a major impact on aircraft operational efficiencies and the ability of airlines to automate practices around the aircraft. But the introduction of 'IT-enabled aircraft' – sometimes also called 'e-enabled' or 'digital aircraft' – enables secure IP communications to and from the aircraft. This is a critical game-changer for the industry. As the first step towards implementing a complete IT infrastructure in the aircraft, it is set to have a major transformational impact on the way airlines operate their aircraft – not only in the cockpit, but also in terms of cabin procedures, aircraft turnaround, maintenance and passenger services. The impact of the IT-enabled aircraft will be all-pervasive, providing the industry with the means to tackle long standing areas of operational inefficiency.

With IP-enablement, we will see new levels of automation and efficiency in cockpits and cabins, enabling crews and passengers to have access to high speed networks and communications. This paves the way for the introduction of new systems, applications and tools on board the aircraft. The reality is already dawning, with the Airbus A380 and new variants of the Boeing 777, as well as with the impending arrival of new types of aircraft such as the Boeing 787 and the Airbus A350.

In the collective mind's eye of the air transport industry, IT has rapidly become a facilitator and catalyst for ongoing aircraft development and optimization efforts. While the potential for innovation may seem boundless, our industry requires a profound convergence at many levels – of standards and regulations, collaborative approaches, infrastructure deployments, and more. Only then can all stakeholders' profit from this digital revolution by enhancing their business efficiencies through a new generation of aircraft operations.

In this chapter, we will first go through an overview of the legacy aircraft communication systems and applications, as well as the existing aircraft IP connectivity solutions, in aircraft operations field up to passenger communications. Service providers' missions and positioning will be outlined. Future communications systems and applications, as envisioned by European (SESAR initiative) and US (FAA NextGen) bodies, will then be presented. Possible scenarios in the role and scope of service providers will also be sketched, with a special focus on the integration/emergence of various service fields, ranging from aircraft operations to passenger-related services. The role of Air navigation Service Providers (ANSPs) and relationship with traditional aeronautical communication service providers will be addressed. Also relationship with communication service providers that are newcomers in the aeronautical market will be analyzed. Then a case study based on AeroMACS (Wimax in Aeronautical band) will be presented: the way the system could be operated, identification of ground providers and interactions between these providers, will be presented.

2. Current and emerging aircraft communication applications and related systems

As defined per ICAO (International Civil Aviation Organization), Datalink services can be categorized as follows:

- Aeronautical administrative communication (AAC). Communication used by aeronautical operating agencies relating to non-safety communications for the business aspects of operating their flights and transport services.
- Aeronautical operational control (AOC). Communication required for the exercise of authority over the initiation, continuation, diversion or termination of flight for safety, regularity and efficiency reasons.
- Air traffic services (ATS). A generic term meaning variously, flight information service, alerting service, air traffic advisory service, air traffic control service (area control service, approach control service or aerodrome control service).
- Air Passenger Communications (APC), defined as services to passengers providing them an access to communications and entertainment means similar to those that can be experienced on ground.

A number of networks and associated communication means can be used by these Datalink services and applications.

3 types of networks are under operation today in the aeronautical world for aircraft – ground communications for cockpit/maintenance/cabin operations:

- ACARS (Aircraft Communications Addressing and Reporting System) used for AAC/AOC/ATC
- ATN (Aeronautical Telecommunications Network) used for ATC
- IP used for AAC/AOC (mostly aircraft IT systems), as well as APC

Depending on the aircraft type, area of operation, and airline, some or all of the above mentioned network technologies are supported. Traditional Aeronautical Datalink Service Providers (DSPs) such as SITA and ARINC provide ACARS and ATN connectivity services. IP air ground connectivity can be provided by traditional DSPs, or by standard telco operators (e.g. 3G operators). These three technologies will in the frame of future ATN (addressed by EU research SANDRA project), migrate/include an IP connectivity solution for ATN.

Depending on the applications to be supported, standard IP connectivity or aeronautical specific ones with specific Service Level Agreements (SLAs)/ Service Level Objective s (SLOs) will be necessary.

In-Flight Entertainment (IFE) and passenger connectivity services are handled nowadays by a variety of subnetworks, especially new aircraft-ground IP links, as well as specific Satcom Inmarsat services. These will be detailed later in the document.

2.1 ACARS

ACARS cockpit data link avionics are installed on approximately 10,000 air transport aircraft and approximately 4,000 business and government aircraft.

ACARS is used by flight operations applications that are hosted in the ACARS avionics unit and is connected to a Multi-Function Control and Display Unit and a cockpit printer that provides input/output to pilots. The ACARS unit is also used as an air-ground router by other airborne systems including the Flight Management system and aircraft system monitoring systems called Digital Flight Data Acquisition Units or Central Maintenance Computers. The ACARS unit communicates with ground networks via various radio systems, always including a VHF radio, and optionally also satellite avionics and/or an HF data radio. Passenger and Cabin application systems can share the use of the satellite avionics if they are installed.

Figure 1 below shows a high level view of end to end ACARS architecture.
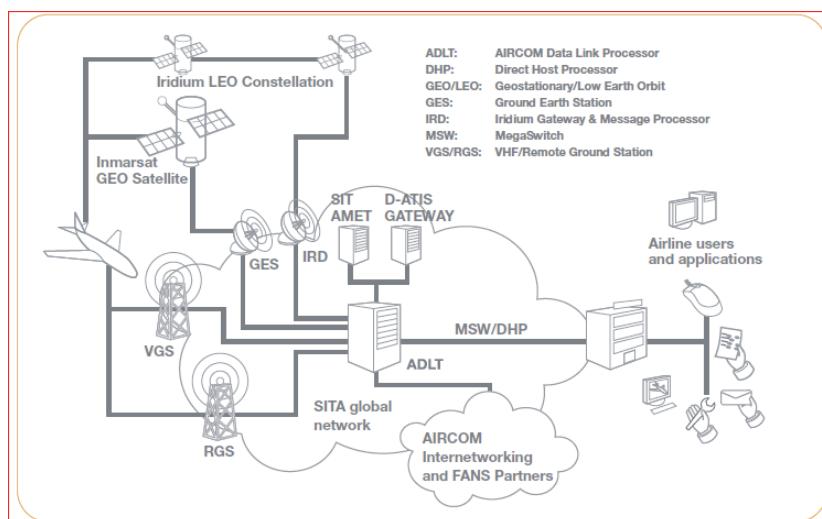
Fig. 1. Overview of SITA ACARS service architecture.

Subnetworks:

The following subnetworks are offered by ACARS service:
- VHF:
- VDL mode 1/A (POA : Plain Old ACARS)
- VDL mode 2 (AOA : ACARDS Over AVLC)
- High Frequency Data Link (HFDL)
- Inmarsat Satcom Data2
- Iridium Short Burst Data (SBD)

VHF and HF subnetworks are operated by the ACARS service providers (DSP), while Inmarsat satellite service is operated by Inmarsat (and, depending on the satellite constellation, also with ground telco partners), and Iridium Satellite service is operated by Iridium. In some countries, the Very High Frequency (VHF)/ VHF Digital Link (VDL) subnetworks are operated by the local ANSP (e.g. China…).

Aircraft communications use of Inmarsat satellite links

Aircraft have been able to carry out voice and data communications via the Inmarsat satellites since around 1990, when these satellites were expanded from their original function of providing services to ships. The number of aircraft equipped to use the Inmarsat aeronautical service today is approximately 2,000 air transport aircraft and another 1,800 business jets or government aircraft. The aircraft using the Inmarsat aeronautical service each month generate a total traffic of approximately 9 million kilobits of ACARS data link messages and 200 thousand minutes of voice calls.

The original Inmarsat Aeronautical service provides two service modes, circuit mode supporting voice communications (or a 2.4kbit/sec modem-to-modem data/fax communications) and packet mode supporting "always-on" data communications.

Aircraft equipped with FANS-1/A avionics (ATC safety communications) use this Inmarsat data link service as the primary means of Future Air Navigation System (FANS) communications in oceanic and remote areas. Aircraft operators use the Inmarsat circuit mode to offer telephony service to passengers and to crew in the cockpit and cabin. Aircraft operators use the Inmarsat packet mode, which provides a data rate of up to 9.6kbit/second, for ACARS communications.

Aircraft data link using HF Radio

The move of aircraft communications from voice to data has motivated some operators of HF radio ground stations to install HFDL computers that enable them to transport ACARS communications.

The vendors of aircraft HF radios have added corresponding capability to support ACARS and it has been installed by a few airlines. The new HF avionics radios can switch between voice and data mode using the same aerial, but they are required to give voice communications precedence over data link, which limits the HFDL availability. A limited number of aircraft are using HF data link and it has been found to provide better availability than HF voice on the routes over the Poles beyond the 80-degree North/South limit of Inmarsat satellite coverage. The HFDL capacity is limited by the frequencies available in the HF band. The allocation of HF frequencies to data link has required a very complex co-ordination process and the system will quickly reach the limits of available capacity. HFDL subnetwork has been also qualified for use for ATS communications.

ACARS Iridium satellite air-ground link

Since 2007 ACARS avionics have begun to be linked to avionics that use the Iridium satellites which fly in low earth orbit and allow avionics to be lighter and less costly. These ACARS messages are being sent in Iridium SBD transmissions. SITA has implemented a gateway between the ACARS service processor and the Iridium SBD server to provide the service via Iridium. Iridium SBD service is being qualified for use for ATS communications.

ACARS over VDL

The air traffic control community defined the ICAO VDL standard to transport ATN air-ground communications but ACARS communications can also use the VDL link. Following discussion of the options for ACARS use of VDL, the Airlines Electronic Engineering Committee (AEEC) Data Link Users Forum in January 1999 adopted as the standard interim architecture "ACARS over AVLC" (AOA).

In the VDL AOA architecture, aircraft use the AEEC 618 protocol over the ICAO VDL standard AVLC link providing 31.5-kilobit per second capacity. Aircraft using VDL AOA obtain increased capacity over the VHF link but can only exchange messages in the same ACARS AEEC 618 formats used over the existing VHF analog link.

ACARS service

The ACARS messaging service is supported by the DSP, and implemented by a set of resilient ACARS message processor(s) (ADLT), with management systems. Internetworking between DSPs is possible, mostly for ATC services.

Who operates and what are the business relationships?

Historical Datalink Service Providers provide ground-ground ATI messaging (TypeB) as well as Aircraft to Ground ACARS messaging. ARINC and SITA have been providing these services for tenth of years now, also with Type A services (aircraft-ground communications). VHF ACARS was the first communication technology developed, later followed by Inmarsat Satellite link. For VHF, several national operators have also been present in this market in a situation of monopoly (China, Japan, etc.). ARINC and SITA have developed internetworking agreements with these local operators to extend their services to these countries.

In early 2000, both ARINC and SITA obtained the right to deploy VHF in the traditional coverage domains of their competitors, i.e. SITA in North America, and ARINC in Europe. This resulted in a service availability improvement (as very often, airlines contract one service provider (primary), as well as a back up, to improve VHF service availability), as well as prices reduction. Only ARINC and SITA provide an ACARS service using satellite communications. ARINC is the only provider of an HF Datalink service for ACARS. SITA relies on Iridium to provide an equivalent coverage (polar routes especially). ACARS supports ATC, AOC, and AAC communications, not IFE/Pax communications, as per regulations.

Ground actors in their strong majority use private network solutions to access ACARS service to aircraft. This is the only reasonable way to meet end to end service level objectives. In a few cases, internet can be used for front end connectivity access to ACARS service, but for non safety / availability application types. It may also be used for downlink only traffic (aircraft to ground). When ATC/AOC come into play, private networking solutions are deemed inevitable by stakeholders to reach adequate security and guaranteed performances. For the same reason, we expect that such communication over IP network will also requires private predictable ground networking for these applications.

2.2 ATN/OSI
2.2.1 ICAO VHF air-ground digital link (VDL) mode 2

The ICAO VHF Digital Link (VDL) Mode 2 standard was developed following the 1990 ICAO Communications Divisional meeting that recognized the value of specifying the use of the Aeronautical VHF channels for data communications. The 1990 ICAO Communications Divisional meeting also reserved the 4 channels 136.900, 136.925, 136.950 and 136.975 MHz for data communications worldwide. Following that meeting, the ICAO Air navigation Commission created the Aeronautical Mobile Communications Panel (AMCP) to develop the VDL standard. The validated VDL Mode 2 standard was presented to the AMCP at its fourth meeting in March 1996, which recommended that it be included in Annex 10. The ICAO member states accepted this recommendation by agreeing to its inclusion in Amendment 72 to Annex 10.

The ICAO VDL Mode 2 standard specifies the use over the VHF link of a D8PSK (Differentially encoded 8-Phase shift Keying) modulation scheme providing a data rate of 31.5 kbits/ second compared to the VHF ACARS rate of 2.4 kbits/second in the same channel width of 25 kHz. The VDL Link Layer protocol specifies for media access control to the VHF channel the same Carrier Sense Multiple Access (CSMA) algorithm as for classic VHF ACARS. However, the VDL CSMA will provide better performance than the VHF ACARS CSMA by using a VHF Data Radio to process the CSMA function. The combination of the VDL D8PSK scheme and its CSMA algorithm makes the link reach saturation at a data load of 10 kilobits per second, compared to the classic VHF ACARS maximum effective link capacity of 300 bits per second.
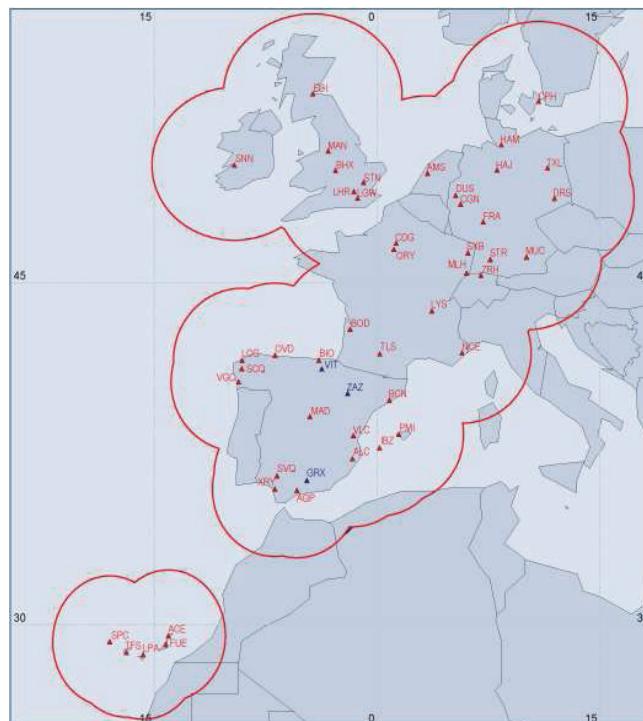
Fig. 2. SITA VDL2 network in Europe.

The "Aviation VHF Link Control" (AVLC) protocol provides a link for the transport of binary data between an aircraft and a ground station. AVLC is a variation of the ISO High Level Data Link Control (HDLC) protocol, designed specifically to handle the use of VHF channels.

ATN (ATN/OSI) End-to-End Protocols

The ATN and data link standards specify protocols using the logic and terminology of the International Organization for standardization (ISO) model for Open systems Interconnection (OSI). The ATN standard covers upper layer protocols used in end systems but this document focuses only on the ATN transport and network layer protocols. The ATN standard specifies ATN applications use of ISO 8073 Connection Oriented Transport Protocol (COTP) over the ISO 8473 Connection Less Network Protocol (CLNP). The COTP protocol provides a message delivery acknowledgement over the CLNP protocol which handles the actual message exchange between the ATN users systems. The ICAO ATN standard specifies a unique addressing scheme to be used in the CLNP protocol which has two formats:

- ANSP systems: ISO county code, city code, terminal identifier
- Airline systems: ICAO airline code, city code, terminal identifier

The ATN CLNP messages are handled by routers that interconnect air-ground (mobile) subnetworks and terrestrial subnetworks. The routers establish a routing information base using ATN routing protocols, primarily the ISO 10747 Interdomain Routing Protocol (IDRP). The ATN routing protocols establish a CLNP routing information base, which is updated as the system establishes subnetwork connections to other ATN systems. Airborne ATN routers maintain a routing information base indicating which connections are available over air-ground data link subnetworks.
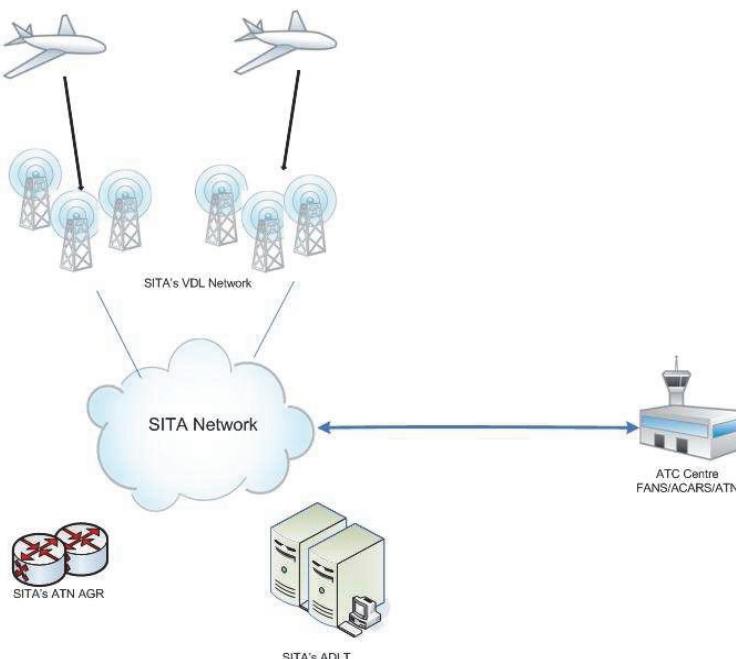
Fig. 3. Overview of SITA ATN architecture.
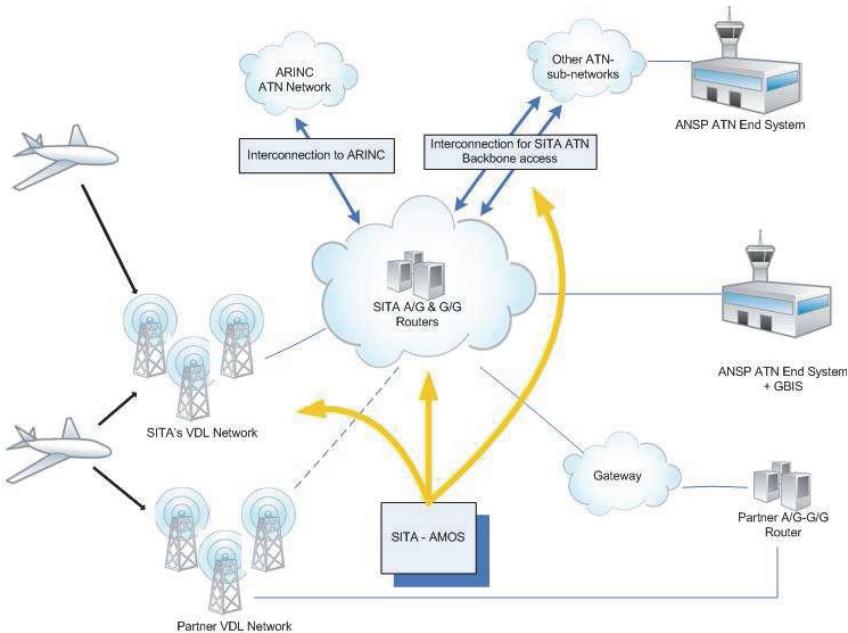
Fig. 4. Overview of SITA ATN architecture – interface with other parties.

It has to be noted that all ATN services can be supported by:
- X25 network infrastructure
- IP infrastructure

Operators are / have migrating to IP (e.g. SITA provides access to AGRs through IP WAN connection (aka IP SNDCF)).

2.3 Emerging IP connectivity for EFB and cabin

In analogy to how modern technology changed and improved modern organisation capabilities, airlines are being convinced that cockpit IT systems will enable them to implement more efficient operations. New cockpit IT systems and applications are likely to rely heavily on IP links for operational purposes in ways analogous to how current systems and processes rely currently on ACARS services.

Comparison of IT equipment using IP in the organization Vs in aircraft communication

Computer have been around for a long time, some of you will remember that hard disks used to be the size of a washing machine and were able to store just a few megabytes of data. Today, people carry in their pockets devices that can hold a thousand times more information. At the same time people were using terminals that could only send and receive 240 characters or 2,400 bits per second as opposed to today's were you can often achieve speed more than 10 mega bits per second using an high speed internet connection.

This is a simple illustration of the technology evolution in the last 25 years. The pace of this evolution as not reduced, on the contrary new technology and new ways to use it evolve at a continuously augmenting rhythm. Each new invention leads to more and more possibilities and also provides the necessary foundations for even more new ideas.

Considering this evolution, the assets used in general Information Technology systems (IT) are subject to frequent changes, shorter duration and amortization. In contrast, technology assets in the aerospace industry typically have very long economic lives and are consequently not usually managed in the same way as general IT assets. Many systems that are delivered in today's aircrafts will remain the same in 20 years. To illustrate the reality of this, we can mention that today's modern aircrafts are equipped with ACARS that is able to transfer 2400 bits per second or about only 10 times more with the newer VHF DataLink Mode 2. This is similar speed to the above 20 years old terminals which cannot be seen anymore in any modern organization. In a recent survey done with New Generation Aircraft (NGA) current and future operators, they have clearly indicated that ACARS will remain an important component of the aircrafts communication infrastructure for the foreseen future.

Although we cannot expect the same lengthy lifespan with new IT systems that gets installed in aircrafts, we can certainly expect that if a technology solution reaches a critical mass of installed based in aircrafts, its corresponding ground infrastructure and associated operational practices, that such solution may certainly last longer in the Air Transport Industry (ATI) then in its counterpart in the general IT market. In the same survey mentioned above, an interesting statement was made about selecting the right technology for a large retro-fit program:

- The time required to accomplish a fleet-wide implementation is longer than the expected life cycle of current solutions. - "We have a vision, but will it be the right one when we have delivered it? We like the technology and we want to use it, but I fear our choices now will be old when the project is fully rolled out".

In consequence, we could anticipate that if airlines and service providers reach a point where massive deployment can be made that the selected solution will last for many years. In addition if as mentioned above this solution life cycle is somehow longer when used in the ATI, that a specific ecosystem will have to exist to maintain it.

What will be this solution? Will this ever happen?

In the next paragraphs we will elaborate on the various conditions and elements that affect the eventual choice and potentials for some sort of industry solution to reach the needed critical mass that will certainly affect the way airlines operates their fleets.

Drivers for new IT systems

In general, key drivers for airlines revolve around attracting and retaining customers; efficient management and operation of their fleets; improving personnel and asset productivity; maintaining safe operations and managing their financial cycles. New cockpit IT systems affect the way aircraft are operated and maintained, so improvements in these areas can result in increased productivity and lower costs

In substance, the main drivers that leads airline to use the capabilities provided by the new technology are the same that drives any other projects:
- Streamlined Processes
- Operational Efficiencies
- Cost Control & Reduction.

The deliveries of new aircrafts such as the A380 and B787 that come equipped with new IT systems are also leading airlines to looks at ways to use these to achieve the above objectives.

Some important barriers need to be considered before projects get implemented:
- The first one, and probably the most important one: The business case that would justify the necessary investment in a major project, its required communication capabilities and necessary enhancements to operational practices is not an easy one. Cost saving and production improvements although perceived as obvious when thinking about the use of modern IT systems and high speed IP communication becomes less obvious when all organizational costs and impacts of a project are added up.
- Typically the IP broadband capabilities of the links available for aircrafts are often justified by the cabin needs, yet it becomes difficult to justify their full use for cockpit operational needs. This lead to the need of cross organizational programs, which are not tradition in a typical airline. So time & effort is required to reach the necessary level of internal co-operation,
- The technology choices have a large impact on costs and no technology solution as yet reached a massive number of aircrafts to be considered as a no-risk choice.
- Security has been high on the radar. However it has been thoroughly addressed by the experienced players, thus security is not seen as a big problem, as long as it is addressed with rigour.

It takes a forward-thinking planner smart enough to envision a way to use the new technology successfully in the design and operation of a brand new aircrafts. Very often successful implementation of new systems and the necessary operational changes would require fleet-wide adoption to get the full benefits. So not only using new systems in new aircrafts conditions the success of a program but consideration to adopt fleet wide systems and processes also become important.

Airlines with multiple fleets will also require assistance to reduce the complexity and differences to deploy maintain and support the various ground systems and communication links required to manage the various IT solution proposed by the air framers and other solution available in the retro-fit market. In the long run, it is expected that airlines will attempt to eliminate dissimilar operational processes and systems across their fleets .This assistance will be provided by their selected suppliers and partners.

As viewed above when choosing a solution, airlines have to consider the solution expected life-time and select the communication technology that will provide the required global geographical coverage at the right performance and right cost. Otherwise they may be eventually at risk of having to support by themselves the entire ecosystem of the selected solution. In such condition, the expected benefits of the new IT systems may not be as profitable as originally expected.

As a consequence of the difficulties to implement new innovative projects using new technology in aircrafts, many airlines that currently operates new generation aircraft that are provisioned for or equipped with wireless IP avionics connected to Cockpit IT systems make limited used of the capabilities at their hubs only an often not at all.

Only very few airlines are currently planning to use new broadband capabilities outside their hub airports or major stations; however, it is expected that the initial delivery of the Boeing 787 now planned for 2012 will bring more opportunities for changes. In addition former manual processes might not even be possible anymore due to turn around constraints and data volumes. Similarly the entry into production of the Airbus A350 (planned in 2013) might also bring significant business incentives to implement new practices relying on broadband IP wireless systems.

Considering the above, airlines have to be careful when considering their choice of partners and suppliers and look for the ones who understand the complexity of airline operation with mix-fleets and that are expected to remain strong players in the ATI for the foreseen future. The current financial status of these organizations as well as their past history would also be a good indication that they can be an adequate choice.

Benefits of using commercially available Off-The-Shelf (COTS) technology

Aircraft undergo severe conditions in their regular journey, with frequent and important changes in temperature and atmospheric pressure. In consequence, systems that must be installed aboard an aircraft need to be designed with the aircrafts particular constrains, needs for security, stability and durability. This lead to careful validation, tests and certification while augmenting the development process complexity.

While using Commercially available Off-The-Shelf (COTS) IT technology and protocols instead of technology that is particularly built for aircrafts may reduce the development cycle, it should be expected that all other particular requirements remains part of the design objectives. As such only marginal saving should be expected to equip and maintain the new IT systems installed aboard the aircrafts.

Certain confusion can be observed in the market with the air framers introduction of COTS IT systems in new generation aircrafts. As the technologies used in aviation applications move from purpose-built to generic, the entry barriers for new entrants have been considerably lowered. Many organizations who historically have not served the ATI industry are now seeking opportunities to extend their market to this field simply based on the fact that the same technology they are familiar with is starting to be used in aircrafts. As seen above COTS IT is only one factor, and not the most important one that must be taken into consideration to adequately serve this particular industry.

The sum of solutions proposed to airlines is so large that no single technology, systems and IP broadband connectivity is yet rising to become an industry standard. While this allows for strong differentiation between airline offers, this also prevents reaching a critical mass that would drive costs and price down for the benefit of everyone. It is important to note that all IT systems installed aboard the aircraft also require their ground counterparts: The back office airline operation systems and at least the global communication infrastructure supporting the chosen aircraft communication method. Massive investment in large and global ground infrastructure will only become possible once a few technology solution raises to become a preferred choice. In the mean time the industry is left with a ground infrastructure that offers disparate coverage, limited to certain region or airports.

Airlines will be looking into achieving the following as listed in Table 1 with the introduction of the new IT systems.

The players

Who can play a role in deploying infrastructure suitable to provide and support new aircraft IT systems and associated ground components?

- Airport can provide local IP broadband connectivity to their hub airlines. This is reality already. This is typically Wi-Fi connectivity from the gate, hangars to the aircraft;
- Airports may also extend their service offering to other airline flying to their facilities. This may work quite well for a few airports and airlines, but will certainly fail attempting to extend the model: Each airline would require entering into a specific and often complex project with many airports to get the necessary connectivity at multiple

outstations. Such task is viewed by many airline as too complex and too costly to be seriously considered on a wide scale;

- Cellular provider with their continuously increased network performance and decreasing prices are already common provider of IP broadband connectivity solution for aircraft. While this model currently works best in the provider local country, the roaming model is less attractive with higher prices. Solutions to these high roaming costs are rising: using device that can support multiple provider SIM cards along with the necessary technical capability that allow choosing the right SIM card base on current location and;
- Global SIM card providers who can negotiate very competitive pricing with multiple providers based on volume and usage projection. These last types of offers, although very recent, seem to be a good model for aircrafts that need to travel in multiple regions. One other issue that may arise with this technology and its communication method is the fact that the cellular networks are usually shared by many users and may suffer from congestion problems. We suspect such problems to rapidly fade away as cellular providers often have the right business justification with the increasing volume of users to invest in enhancement of their networks. Global SIM card provider may also benefit from being already able to offer network connectivity using dedicated circuit which allow offering services with the level of performance and stability suitable for airline operational activities while not suffering from the congestion problems observed with local cellular providers;
- Global Satellite provider such as Iridium and Inmarsat offer IP broadband service through distribution partner mainly for in-flight connectivity. Each major provider has offers that competes and have gained some market share.
- Global datalink service providers already servicing the ATI, mainly SITA and ARINC are developing solution to address the IT systems and connectivity needs of the new aircrafts systems.
- New entrant in the ATI industry. The introduction of cockpit IT brings new opportunities to companies currently absent in the aircraft communications and connectivity market but who may have experience integrating and offering ground-based broadband and IP solutions. Many are looking to complement their existing portfolios and move into aircraft IT systems and connectivity. In addition the availability and commonality of IP-based network connectivity through increasingly accessible satellite operators and with the Internet reduces necessary investments in some infrastructures to offer solution that may win some of the aircraft IT and connectivity business.

As of now the industry in general, benefits from a mass of expert in some of the common aspect of the systems that gets installed in aircraft and in particular, the IP protocol used by these systems. Considering the difference in life cycles of the technology assets that gets installed in the aircraft and the IT industry in general, this mass of expertise might only be a temporary situation lasting only until IT systems reach a new level of sophistication. We may reach a time where the general perception is that aircraft IT systems are archaic, and only perceived as fit for this unique vertical market, very similar to today's situation with legacy aircraft systems. In consequence, airlines when selecting their aircraft IT suppliers must be careful to consider their overall ATI expertise and their real chance to last in this particular industry.

Expected Benefit	Supporting Features
Cost savings and operational improvements through better and greater access to aircraft and post-flight data while on ground	Wireless broadband access to aircraft and flight data upon landing during turnarounds and layovers Use of low cost wireless technology to transfer large amounts of non-critical data (trending, manuals, etc.)
Flight operational efficiencies achieved through access to most recent data prior to departure	Ability to send flight crews latest information while at the gate.
Process automation and labor cost reductions in low value activities	Wireless data transfers represent a vast improvement over manual data retrieval and documentation updating processes Resources and time can be shifted to higher value, more productive activities
Aircraft as a flying node in airline IT infrastructure	Seamless integration to back-office and 3rd party systems and processes to enable differentiation, cost saving and revenue generating initiatives.
Secure connectivity	Ability to offer a flexible yet secure environment specifically tailored to address the needs of aircraft communications over IP
Productivity improvements	Aircraft use of Electronic Flight Bags (EFBs) using IP communications will allow them to reduce the amount of paper needing to be provided to pilots and processed manually after the end of flights enabling airlines to save money
EFB usage reduced operational costs/Savings	Saving are then based on crew's productivity gains and better flight decision due to more accurate and up-to-date-information as well as a possible reduction in weight (no more paper to carry).
Competitive advantage	Airline installation of new IT systems, IP communications and operational practices will enable them to be more efficient than their competitors

Table 1. Overview of benefits expected from introductyion of new IT systems.

It should also be expected, in the next few years, a high rate of adjustments in this market with strong consolidation of the various offers and many partnerships between software shops, avionics vendors, service providers and others.

Who want to play in this game?

In general, multiple technologies have proven their technical validity to provide high speed IP connection to aircrafts. Most should be capable of supporting both cabin and cockpit communications (even if they currently don't):
1.   Based on ground wireless network:
   - Generic 3G cellular network for ground communication

- Wi-fi with the Gatelink solutions;
- Cellular network directed toward aircraft such as the Aircell solution operating in the US.
- Wimax that is being introduced by a number of vendors
2. High speed geostationary satellites:
   - Ku-Band
   - Inmarsat new I-4 constellation of satellites
3. Low orbital altitude flying moving satellites
   - Iridium constellation of 66 satellites that can support data service up to 128Kbps
   - Iridium next generation satellite network, NEXT, planned for 2014/15

Each technology has its merits and limitations. As such it is expected that most will be in the market for a number of years to come. What is less certain concerns the right commercial approach to develop and gain market share. Regulatory aspects are also important, and they will also affect the adoption of one technology over another for cockpit communication.

Today no standard approach has been adopted by airlines to implement new practices and infrastructure to accommodate broadband aircraft communication capabilities. Essentially, each implementation of supporting infrastructure has been unique. The only common IP broadband technology installed in today's major airframers New Generation Aircraft (NGA) is the availability of a Terminal Wireless Lan Unit (TWLU) capable of wireless IEEE 802.11 connectivity from and to cockpit systems. Cellular connectivity is also widely used, but is not generally connected to cockpit systems other than to the Quick Access Recorder (QAR).

EFBs are the main type of cockpit IT systems in use today. They require a level of resilience above that needed by the non-critical applications, but the data exchanges, at least initially, still can be limited to hubs and main stations; however the increasing complexity of EFB applications will also make the IP wireless links (whether in the hubs in flight or at out stations) and overall connectivity to airlines own networks much more critical for airlines operations.

In addition to the normal use of ACARS for in-flight AOC and ATS communications, NGAs will use more and more IP-wireless links not only to refresh the contents of their EFBs and IFEs or to download massive amounts of flight and aircraft-related information, but also to upload critical software parts and to access third parties' networks while on the ground. The availability and coverage of IP wireless links will shortly become paramount for airlines operating these aircraft types. As of now most airlines have selected to use the wireless IP broadband link at their hub only but the increase fleet size and requirement to cut costs and increase productivity, is starting to trigger projects that aims to use this same capability at out-stations. As such service providers will shortly be under considerable pressure to provide IP-based services and coverage to facilitate the dispatch-ability of the aircrafts regardless of where they fly.

The upcoming Boeing 787 will bring a marked departure from other NGA already delivered. In fact current Boeing 777 and Airbus A380 operations do not imperatively require IP broadband wireless connectivity. Data transfer can be deferred to hub-only operations.

As mentioned before, the 787 fleets with its increased complex IT systems will bring more opportunities for changes, since the volume, frequency, and criticality of exchanges of operational information between the aircraft and ground systems is expected to be higher than for older aircraft. It is expected that without the use of wireless links such as GateLink,

to transfer 787 data prior or after every flight that airlines may run into operational inefficiencies. Consequently early 787 customers are expected to lead the way in their adoption of new operational practices and systems surrounding aircraft connectivity, and they form the primary initiators and requester of changes that may lead to a wider scale industry adoption of standard solutions than what we have seen up to now. The same can be said about the next up-coming new fleets such as the Airbus A350.

As an example to illustrate this need of increase data exchange for new aircraft is the Quick Access recorder (QAR) data, known in the 787 as continuous parameter logging, or Continuous Parameter Logging (CPL) which could produce up to 100 MB per flight. This can take a considerable amount of time to download manually, and could be lost if the transfer is not completed before the system memory is exhausted and overwrites the earliest data. The same may apply to the engine health monitoring (EHM) data.

All other aircraft types (including 777s and A380s) can now be handled using legacy services and practices, so there are presently limited incentive or urgency for undertaking the significant investments to install or use new wireless technology even if the IT systems installed in these aircraft allows such use.

In resume expectation are that early Boeing 787 operators may lead the way in their adoption of new operational practices and systems surrounding aircraft connectivity.

3. Future communications systems and applications

The future SESAR ATM concept demands datalink services supporting features such as 4D trajectory management, ASAS separation, automation, and SWIM. A reliable and efficient communication infrastructure will have to serve all airspace users in all types of airspace and phases of flight, providing the appropriate Quality of Service needed by the most demanding applications. The mobile part of this infrastructure will be based on a multilink approach, composed of three different subnetworks:

- LDACS: A ground-based line of sight datalink as the main system in continental airspace and supporting Air/Ground services and possibly Air/Air services, offering a high Quality of Service which will be necessary in the high density areas; two systems are under consideration (LDACS 1 and 2) with the objective to select one for implementation. Both operate in the L-Band and are based on modern and efficient protocols;

- Satellite: A satellite based system providing the required capacity and Quality of Service to serve oceanic airspace whilst complementing ground-based continental datalink as a way of improving the total availability. The system is being defined in close cooperation with the European Space Agency. The type of satellite constellation to be used (dedicated or commercial) is still under consideration;

- AeroMACS: A system dedicated to airport operations, based on mobile Wimax 802.16e, providing a broadband capacity to support the exchanges of a significant amount of information such as the uploading of databases or maps in the aircraft.

In addition, and to allow in the medium term interoperability with military operations, a gateway is being defined to interconnect the ATM system and the military link 16 system.

Several research programmes have been launched to define, develop, and validate these new solutions, and prepare the Aeronautical community to transition to these new access networks. These activities are handled within SESAR programme. The SANDRA project also takes into account the integration aspects of these new solutions, and the networking environment (IPv6 will be introduced in place of IPv4).

Figure 5 gives an example of various networks and consequently operators that could be involved in future ATS/AOC communications.
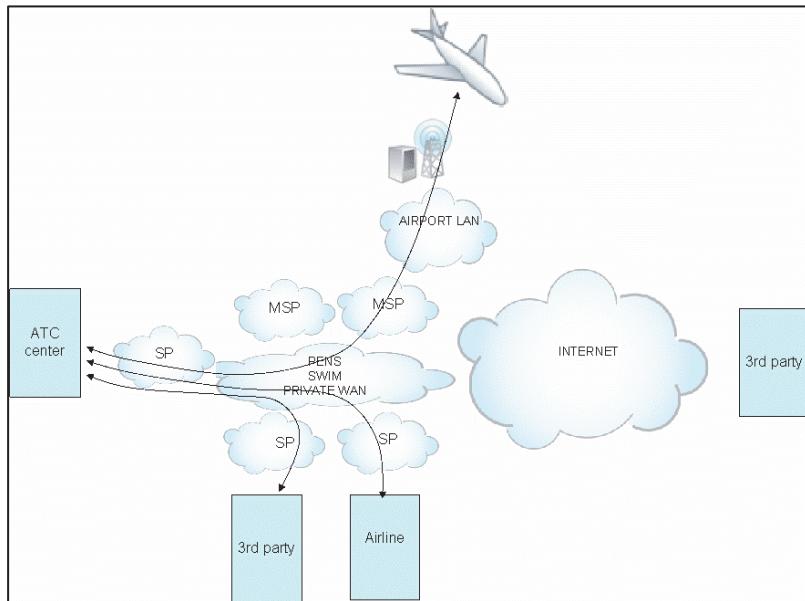
Fig. 5. Example of networks and actors that could be involved in future ATS/AOC communications.

One major difference with IPv4/existing IP connectivity services is that mobility management will probably be provided as a service. Mobile IP being part of the basic scope of future ATS communications. Mobility Service Providers (MSPs) will thus be needed. We can imagine of course having different MSPs between the APC domain and other domains. The service provider on the APC side may even not be aero-specific. However, when we compare these new schemes with the legacy ones (will be detailed in the following chapters), the main actors types remain.

4. Analysis of service providers' roles and business relationship

4.1 Now (ACARS, ATN, IP)
The ACARS business relationship can be modelled as shown in the diagram of Figure 6. With a limited number of organizations dealing in this market, the model is very simple. The actors identified are:

- Users: Airline (Aircraft), ANSPs (Ground ATC End systems), 3rd parties for AOC/AAC
- Global DSPs, providing ACARS service to aircraft and ground access to Aircraft using ACARS service. This is globally limited to two organisations: ARINC and SITA. Global DSP operate also their own VHF/VDL 'almost' worldwide network.
- Local VHF and/or VDL mode 2 operators, providing VHF/VDL2 ACARS service to aircraft, and ground access to global DSPs – a few ANSPs operate their own VDL/VHF network – the trend is however to outsource the network service
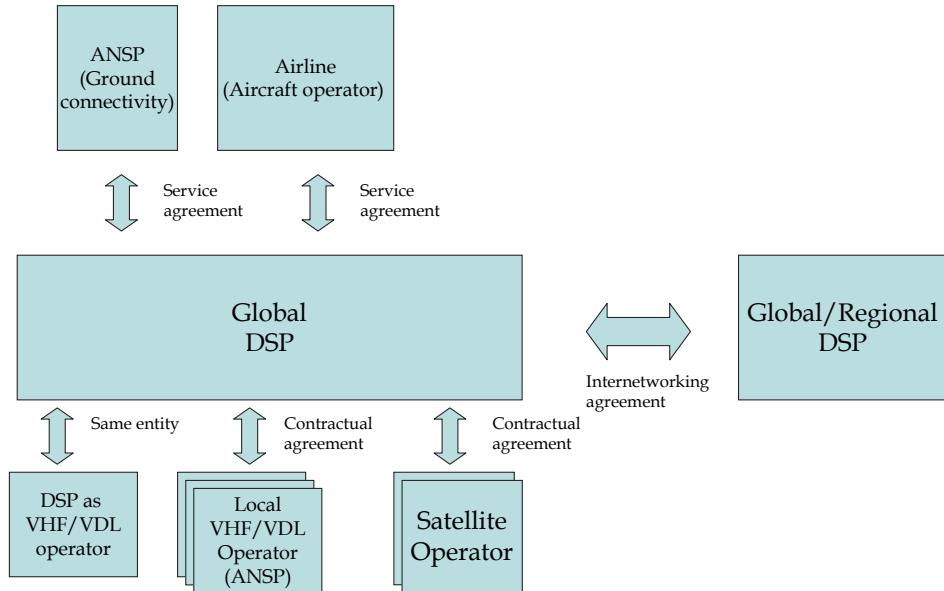
Fig. 6. Illustration of actors' relationship for ACARS.

4.2 Focus on existing roles and actors in ATN/OSI

The ATN/VDL2 business relationship can also be simply modelled with a limited number of organizations:

- Users: Airline (Aircraft), ANSPs (Ground ATC End systems)
- ATN operators, providing ATN networking service
- VDL mode 2 operators, providing VDL2 access network and connectivity to Ground ATN network

However, it has to be noted that ANSPs either purchase the VDL2 service to 'global operators' such as SITA and ARINC, or operate the VDL2 service themselves and allow global DSPs to manage the AOC traffic. Even if the overall trend is to outsource such service to partners able to sustain the liability and SLA constraints of safety and dispatch oriented services, these two models will likely be found for future communication means (IP).

4.3 Focus on new roles with the introduction of new IT systems

The new business relationships become more complex in the new aircraft IT world with many more players:

- Users: Airline (Aircraft), ANSPs (Ground ATC End systems)
- ATN operators, providing ATN networking service
- VDL mode 2 operators, providing VDL2 access network and connectivity to Ground ATN network
- Global DSPs, providing ACARS service
- Global IP communication service provider
- Regional IP communication service provider
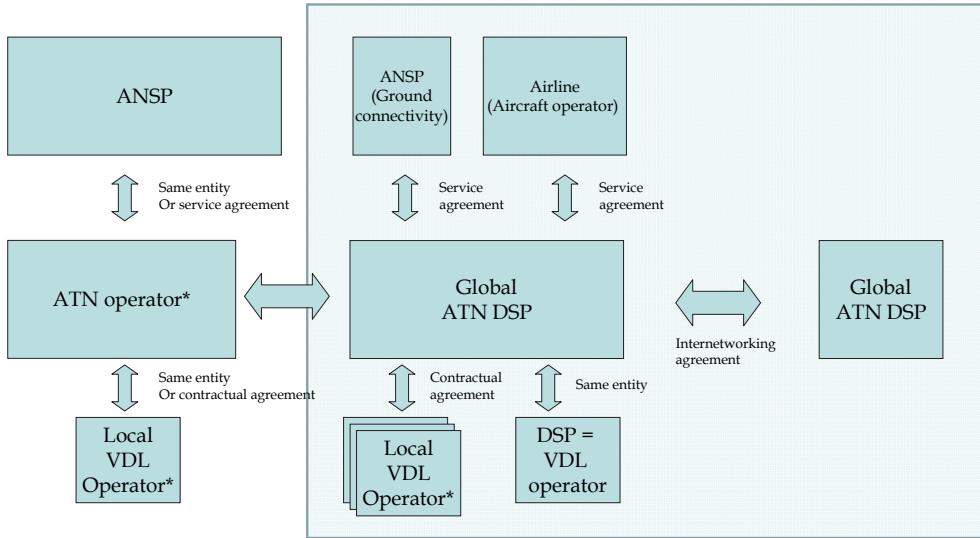
Fig. 7. Illustration of actors' relationship for ATN/OSI over VDL mode 2.

- Local IP communication service provider
- Access network operator (e.g. Inmarsat for SwiftBroadband)
- Solution integrator
- Avionic vendor who now offer multiple IT solutions, communication services and back-office solutions.
- Airframers IT solutions
- Airports networking services
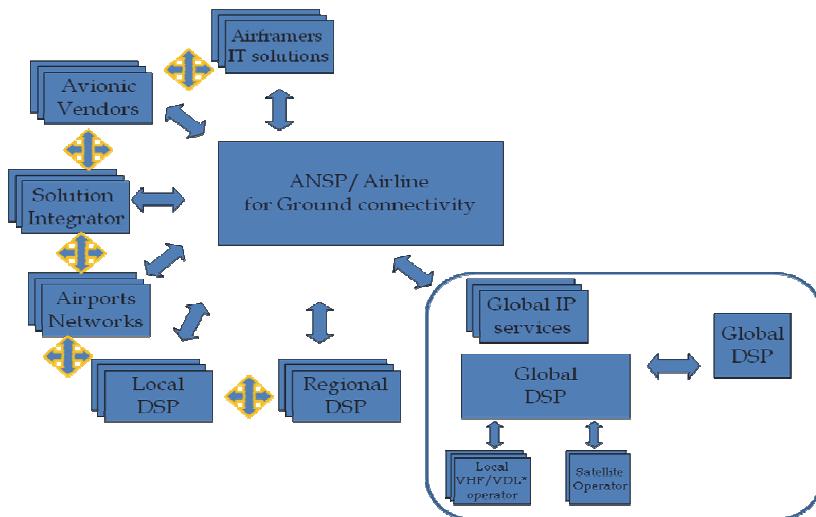
Fig. 8. Identification of main actors for new IT systems operations.

If we take the example of SwiftBroadband, several major actors are involved
- ASP (application service provider): provides the application using SBB Satcom services
- SP (Service Provider): resells airtime and services to airlines; may activate SBB channel if delegated from DP; may have own APN (Access Point Name: network node on ground for traffic routing), if agreed with DP
- DP (Distribution Partner): SBB channel activation (one SIM card per channel); resells airtime to Service Provider; may directly retail airtime to airline (e.g. OnAir) - DP is linked to SIM card, thus potentially one DP per SBB channel
- Satellite / Swiftbroadband service Operator (Inmarsat)

The actors listed above, specific to SwiftBroadband are Inmarsat and the DP. The other ones can easily perform horizontal integration (with other access networks). A given partnet can perform vertical integration (act as a DP and ASP/SP). All combinations are possible, several DPs per aircraft, several SPs per SIMCard, etc. However, it is of course strongly advised, in order to make the system manageable, to minimize the number of actors and rely on key players.
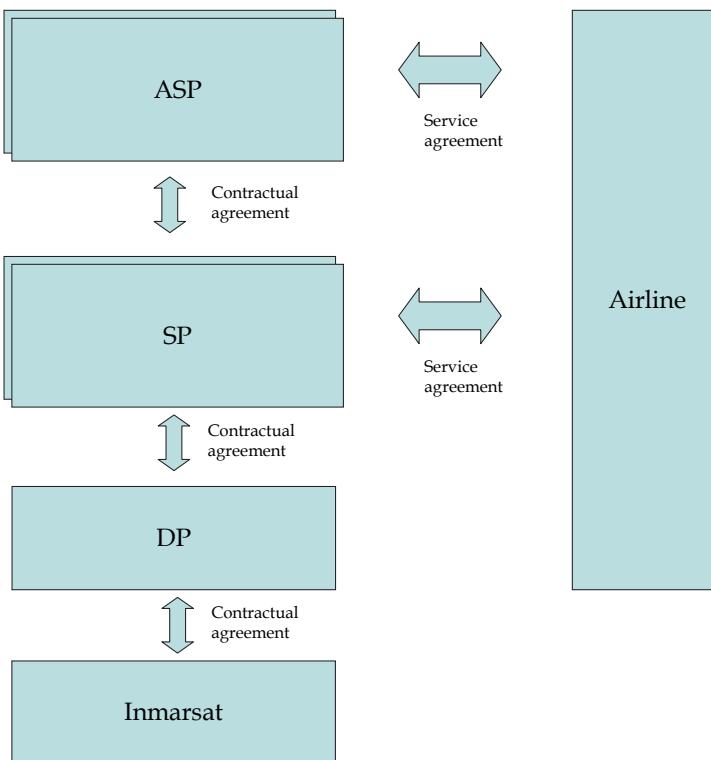
Fig. 9. Actors and relationship for SwiftBroadband

4.4 What could be the winning combinations after cards have been shuffled again?
4.4.1 Key technologies and integration on aircraft

Some of the key enabler to an eventual global successful aircraft connectivity solution, is the availability of adequate aircraft-ground connectivity technologies, at the right performances, with worldwide availability, at the right price and ability to integrate these technologies in a large retro-fit program, at minimum cost and reduced aircraft down time. Security of the solution is also imperative.

Such solutions must then:

- Provide sufficient coverage/availability (regional, airports, etc.)
- Be implementable at minimum cost on aircraft or be provided as part of wider system

A number of proprietary solutions such as Aircell, Teledyne WGL/QAR exist and reached an interesting level of success.

Adding new technology, new providers and new application creates an environment that is becoming exponentially complex. At the end, airlines being successful will definitely need to be able to make the right choices, reduce their risks and be carefully to limit their investment to solutions that will last. One of the factors that enables meeting these objectives and constrains is to share those risks and investments with industry partners. Another key element to choose adequate partners is the ubiquity of the solution they propose. As airlines fly everywhere, the solution chosen must be available globally. Solutions that remain only available in certain geographic location may certainly last in a specific market, but have not the potentials to become industry standards.

Regulatory aspects

A number of regulatory requirements and actors come into action when we talk about aircraft communications.

- Operating an access network generally imposes the use of radio licenses
- Dealing with ATS communications implies to interact with national ANSPs as customers or as providers/partners
- And many others

The ability to deal with such entities is a prerequisite to global service provisioning. Of course, but this is a special item, an aircraft embedded system needs to be certified at appropriate assurance level

Vertical and Horizontal integration

It is interesting to focus on the positioning of the success players in datalink history and see how things are evolving. Horizontal integration at access network and network level is compulsory to provide consistent services, and add value to the fact of integrating multiple dissimilar access networks (with their incumbent complexity due to the multiplicity of operators).

This is what happened in ACARS and ATN. Traditional DSPs started developing with vertically integrated solution (VHF – ACARS – some application services) to horizontal integration to make the services available worldwide (VHF operated by DSP and other access network services outsourced (Inmarsat). Traditional DSPs have positioned themselves more in the Cockpit communications domains than on cabin domain.

Another important factor here is the existence of historic operators/compulsory operators that are imposed by local regulations. The ability to deal with such entities is a prerequisite to global service provisioning.

We could define vertical integration by the fact of acquiring the ability to control parts or all the actors/functions needed to provide an overall service, i.e. providing at the same time different levels of service (access, network, application, etc.), while horizontal integration could be defined as the fact of acquiring the ability to widen geographically or in terms of market target (e.g. cabin versus cockpit) a given service.

Figure 10 illustrates this concept. It is interesting to see that, depending on the market segment (ATS/AOC legacy,…), and the service level, the position of existing bridges (similar products that can satisfy upper services) can vary. For example, it is obvious to notice that ATS/AOC and EFB will likely call similar skills and services (IT integration, data production and publishing), while communication means between EFB and cabin could be shared (e.g. SBB,…).
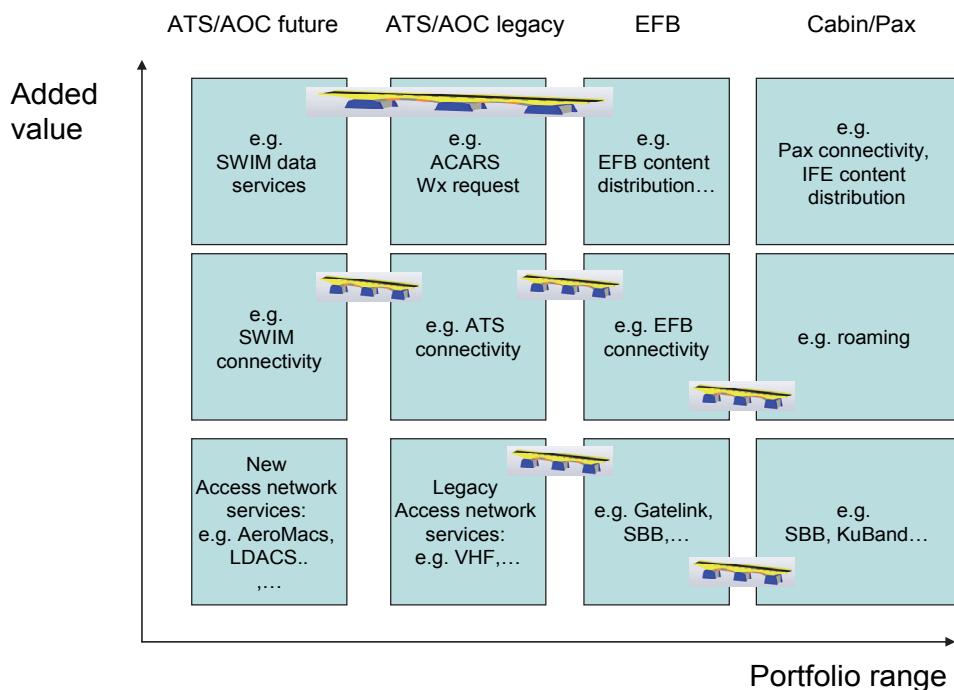

Fig. 10. Illustration of Vertical and horizontal integration.

Current and future NGA operator's views

The recent survey with NGA operator mentioned earlier includes many indications that confirms and support many of the information given in this chapter. Here is a small overview of what some of the key players in airline operation are mutually saying about using new IT technologies in aircraft operation:

- Understanding the complexity and inter-relationship that will shape the future of aircraft operation result in a long learning curve.
- The requirement for cross organisational collaboration is viewed as an essential element for a successful program to implement new technologies for aircraft operation.
- Many delays are caused by the needs for common understanding and alignment of the multiple parties involved, including regulatory authority, airframers and standard bodies.
- Obsolescence of chosen technology in contrast to the life time of fleet-wide implementations is a major concern an often a road block to making technological choices.
- There is a tendency to make incremental steps forward as the solutions and industry vision evolves.
- The search for real business value leading to a successful business case is a difficult task in the current context.

5. Conclusion

Airlines expect to be able to meet their short and long term business objectives using the new IT technology available in the aircraft cockpit. No specific solution, IP broadband communication method or technology as yet rises to become the industry standard necessary to limit the risk associated with large deployment project. This is the case for both airlines and service providers. Legacy system and communication technology installed in today's aircraft will remain pertinent for the foreseen future and need to be integrated in the offered IT solutions. The complexity associated with the installation and operation of new IT systems is continuously rising. Absolute confidence in watertight security of the new systems and communication links must be achieved.

As the technologies used in aviation applications move from purpose-built to generic, the entry barriers for new entrants have been considerably lowered. Consequently the complexity and diversity of the solutions and required inter-relationship of the industry players is considerably augmented. Providers have to be carefully chosen by airlines based on their offer of valuable and compelling service that can assist them to make the most efficient use of their modern aircraft without compromising their operational flexibility or security. Solutions that can be built to take into consideration the various technology choices, requirements for global availability, the typical aircrafts projects life-time, integration with legacy systems and needs for common approaches will certainly have a better chance to be successful in the long term.

As we have seen above, from the customer's prospective, dealing with major partners providing horizontally integrated solutions, **especially at access network level**, will likely be the way to go, providing that the investment stays reasonable and offers an interesting return on investment scheme. Of course, horizontal integration should target the pertinent access network technologies (efficient, reasonable cost) that can be deployed on aircraft at

reasonable cost and cycle. Each customer's case is specific, so it is of course too simplistic to summarize it this way, but this trend might well prove to be true.

It will be several years before new cockpit technology deliver on its expected benefits to reduce overall fleets operational cost and improve productivity, but we are clearly heading in that direction.

Possible customers' perspective

We do not take much risk if we say that customers seek

- Low cost
- SLA/SLO and adequate performances
- Globally available service – at least on strategic routes or locations
- If possible, end to end managed service
- And now, proper integration in their operations process (integration in their IT environment or hosted IT environment)

For some specific discriminating services towards their competition, some airlines may be willing to invest to offer unique services to attract new passengers, for example in the domain of aircraft passenger services.

We could conclude from this chapter that, in order to make future connectivity services a success, the airlines will seek for service providers

- Horizontally integrated
    - Multiple radio access networks and ground networks
- Vertically integrated
    - Application services
    - Up to the IT infrastructures

And of course there is a STRONG

- Need for competition
    - Standardization
    - Multiple players

This is a general conclusion, and, as said before, airlines needs need to be studied on a case by case basis, but we tried here to give general trends that will hopefully help the reader have a wider view of the situation.

6. Appendix – case study: AeroMACS

This section aims at introducing the various possible actors in AeroMACS connectivity service and identifies their possible contractual relationship.

The following applications have been identified as target by RTCA SC223 and/or Eurocae WG82:

- Fixed users (RTCA only)
    - Airport LAN extensions
        - o Unique equipment (terminals, cameras,…)
        - o Or Multiple equipment behind Mobile System (MS)
    - ANSP managed equipment
        - o Integration of RNAV systems, radar… into ANSP network
- Mobile users

- Airport trucks (catering, maintenance, fuel….)
- Luggage terminals,….
- Single users (= user terminal) belonging to different networks or LANs (!) (airport, MRO, catering, fuel……)
- Support for VoIP (RTCA only)
- Aircraft
  - ATC, AOC,direct operational impact / safety impact
  - AAC applications no direct operational impact
  - End to end (to airline and ANSPs) communication but also potentially local communications (cache servers)
  - Need to segregate on aircraft at minimum between various users domain (avionics (ATC, AOC) – IT domain (AOC, AAC) – Pax domain)

6.1 Actors and possible business / contractual relationship

WiMax forum (WMF) Network architecture group has identified the following typical business relationship for WiMax as shown in Figure 11.

Fig. 11. WiMax forum identified actors and relationship.

In the aeronautical environment, the following actors have been identified:
- NAP
  - Specialized Airport operator
  - Traditional DSP (ARINC, SITA, AVICOM,….)
  - ANSP
- V-NSP
  - Traditional DSP (ARINC, SITA, AVICOM,….)
  - ANSP
  - Local operator
  - Specialized Airport operator
- H-NSP
  - Traditional DSP (ARINC, SITA, AVICOM,….)
  - Others

In Figure 12, the reference contractual/business relation ship between various actors in the aeronautical environment could be as illustrated.
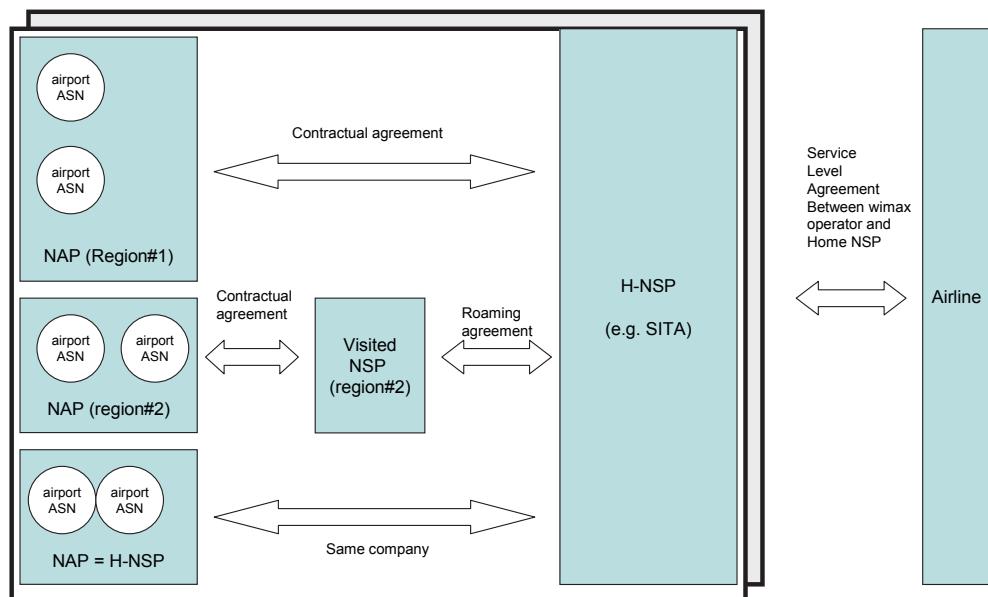
Fig. 12. Actors and possible relationship for Aeromacs.

6.2 Network deployment use cases

The following deployment use cases have been identified by WMF (reference).:

- A.3.1 NAP Sharing by Multiple NSPs
- A.3.2 Single NSP Providing Access Through Multiple NAPs
- A.3.3 Greenfield WiMAX NAP+NSP
- A.3.4 Greenfield WiMAX NAP+NSP with NAP Sharing
- A.3.5 Greenfield WiMAX NAP+NSP Providing Roaming
- A.3.6 Visited NSP Providing WiMAX Services
- A.3.7 Home NSP Providing WiMAX Services

The deployment use cases identified above are instantiated for aeronautical environment as follows – please note that multiple use cases will be supported at the same time. Especially, an aircraft will need to be able to interact with various use cases, depending of its location at time T.

Open points to be discussed:

- Multiple NAPs will be available in a given airport and one used at a given time by an aircraft
- An aircraft may contract several H-NSPs and selection will be done in real time.
- Need for NAPs to advertise supported NSPs (H-NSP??) and the aircraft will then select the preferred H-NSP

6.3 RF deployment use cases

Several radio / NAP deployments are possible:

1. Single NAP – single radio channel
   - Use service flows / QoS to distinguish between application types (aircraft)
   - All NSPs advertise on this channel
2. Single NAP – Multiple radio channels
   - 1 channel for aircraft communications
   - All NSPs advertise on this channel
   - 1 or several channel for fixed and mobile users
3. Single NAP – Multiple radio channels
   - 1 channel for safety communications (ATC/AOC) for aircraft and mobiles
   - 1 channel for non safety (AAC), including mobiles
   - 1 channel for fixed users (RTCA use case)
   - Implies 2 radios on aircraft. This solution is probably not adequate due to aircraft systems complexity
4. Multiple NAPs – Multiple radio channels (NAP+NSP solution)
   - 1 several channels for fixed and mobile users

Various constraints need to be taken into account to determine the appropriate solution(s):

- Channels allocation scheme selected by various states/countries
- Limit configuration changes on aircraft side
- Segregation between non safety/dispatch impacting traffic and non safety / non dispatch impacting traffic
- An aircraft will need to interoperate with any AeroMACS infrastructure, while ground mobiles and fixed users may be tailored to specific ground infrastructures
- Fixed infrastructures support safety traffic may need to be severely segregated from mobile infrastructures
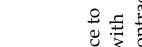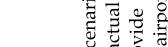
Use case	Needed?	Architecture overview	Example
A.3.1 NAP Sharing by Multiple NSPs	yes		e.g. airport telco operator deploys and provides service to ARINC, SITA.. As H-NSP who manage relationship with airports on behalf of the airlines. Some airlines may contract one H-NSP, others other H-NSPs.
A.3.2 Single NSP Providing Access Through Multiple NAPs	yes		e.g if NAP= different entities than H-NSP, then this scenario is needed. Typically, DSP (H-NSP) may establish contractual agreement with various airport telco operators to provide overall connectivity to airlines. ASNs are in different airports.
A.3.3 Greenfield WiMAX NAP+NSP	yes		e.g. H-NSP (SITA, ARINC..) can also deploy itself at airports (or some airports).Several H-NSP may deploy on different channels in same airport.

Table 2. Network deployment use cases (part I).

Use case	Needed?	Architecture overview	Example
A.3.4 Greenfield WiMAX NAP+NSP with NAP Sharing	no		e.g. H-NSP (SITA, ARINC..) can also deploy itself at airports (or some airports) and provide access
A.3.5 Greenfield WiMAX NAP+NSP Providing Roaming	future		e.g. When Route optimization is supported, roaming could be provided over future ATN (not really internet as shown on the schematics)

Table 2. Network deployment use cases (part II).
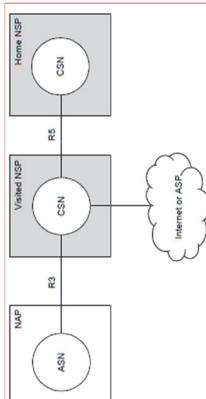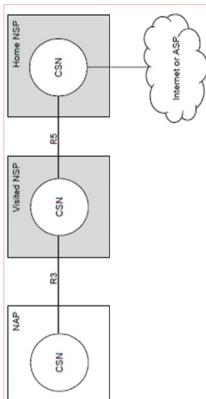
Use case	Needed?	Architecture overview	Example
A.3.6 Visited NSP Providing WiMAX Services	future		e.g. When Route optimization is supported, roaming could be provided over future ATN (not really internet as shown on the schematics)
A.3.7 Home NSP Providing WiMAX Services	yes in some marginal cases?		e.g. in some cases there may be a need to go through a V-NSP (e.g. in some states such as China, you may be compelled to go through the official NSP)

Table 2. Network deployment use cases (part III).

6.4 Conclusion on the AeroMACS case study

We have seen that many deployment configurations were possible for AeroMACS connectivity service. The deployment / contractual relationships and associated actors are still under discussion in various standardization bodies. However, we can see that the number of actors and the variety of situations will likely drive the need, as seen in legacy/emerging connectivity services, for overall mobility service providers, aka Global DSPs, to hide the complexity of the connectivity service schemes to be establish to provide a global, seamless, simple and efficient service to the various service customers.

7. Acknowledgement

8. References

WiMax Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 3 – Informative Annex] Release 1.0 Version 4

ICAO ANNEX 10 to the Convention on International Civil Aviation Aeronautical Telecommunications (Volumes I, II, III, IV and V)

SITA AIRCOM new generation services Positioning Paper – White Paper – 2010

SITA - The "Digital aircraft" –heralding a new generation of aircraft operations – New Frontiers paper – 2010

WiMax Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 0] Release 10 Version 4

(WMF-T32-001-R010v04_Network-Stage2-Part0_0)

WiMax Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 1] Release 10 Version 4

WMF-T32-002-R010v04_Network-Stage2-Part1_2_0

WiMax Forum Network Architecture (Stage 2: Architecture Tenets, Reference Model and Reference Points) [Part 2] Release 10 Version 4

WMF-T32-003-R010v04_Network-Stage2-Part2

Part 2

Future Aeronautical Network Aspects

SOA-Based Aeronautical Service Integration

Yifang Liu[1], Yongqiang Cheng[1], Yim Fun Hu[1],
Prashant Pillai[1] and Vincenzo Esposito[2]
[1]School of Engineering, Design and Technology, University of Bradford
[2]SELEX Sistemi Integrati S.p.A.
[1]United Kingdom
[2]Italy

1. Introduction

Over the past two decades, the air transport industry has experienced continuous growth. The demand for passenger air traffic is forecast to double the current level by about 2025 (European Organisation for the Safety of Air Navigation [EUROCONTROL], 2006). Small-to-medium sized low cost airlines in Europe such as EasyJet and Ryanair have observed a considerable percentage of passenger increase between 2008 and 2009 due to the growth in the number of regional airports and more choices offered on international destinations (EasyJet & Ryanair, 2009). To accommodate such growth and changes in new flight patterns and strategies, it is of paramount importance to ensure air transport communication systems around the globe be integrated to enable efficient air-to-ground and ground-to-ground communications for global air traffic management and coordination.

Traditional approaches for aeronautical system integration in the past impose a high level of system dependencies; a fixed connection is required to be set up every time a new application is added. Therefore, aviation companies are facing continuous investment increase every time a new connection is established. This situation discourages enterprises from fulfilling grater business values by adding interior constraints; it restricts the number of applications and services that can be integrated into the existing IT infrastructure. In safety-critical systems in the aeronautical context, overloaded complex system structure will increase the chances of operational failures and jeopardise passenger safety. Therefore, it is important to devise a suitable architecture which minimises system dependencies and allows new applications to be integrated easily with the lowest IT maintenance budget. A layer-extensible blueprint in a Service-Oriented Architecture (SOA) is considered as a solution in this case for the integration of future aeronautical communication systems. The proposed framework should allow consistent data capturing and sharing among all end users who are involved in the global aircraft operations in the 2020 timeframe and beyond.

In recent years, the SESAR SWIM (Single European Sky Air Traffic Management Research System Wide Information Management) concept has reflected the emerging needs and willingness of Air Traffic Management (ATM) organisations in transforming proprietary ATM systems into a standardised and interoperable information pool in the pan-European aeronautical network. As the challenge still exists where the ATM stakeholders today do not want to deal with the complexity of the lower communication layers, SOA is considered as

one of the most effective emerging technologies to provide a scalable, flexible and interoperable system framework, according the adoption of the SOA paradigm in the global SWIM studies (Houdebert & Ayral, 2010; Luckenbaugh et al., 2007). However, SWIM focuses on ground-to-ground aeronautical services.

In extending the SWIM ideology to an airborne context, the EU FP7 project SANDRA (Seamless Aeronautical Networking through integration of Data-Links, Radios and Antennas) continues with the SOA notion in air-to-ground information exchange, service composition and integration to provide a complete and coherent set of communication services for Next Generation global Air Traffic Management.

Building on the investigation and analysis of the existing industrial programmes targeting aeronautical service integration, this chapter provides an introduction of the SOA-based future avionic systems. Section two outlines the middleware concept, the service-oriented architecture, its implementation technologies and the use of SOA design solutions forming an integrated ATM framework. Section three summarises the SOA-based future aeronautical communication referring to the SESAR and FAA (Federal Aviation Administration) SWIM approaches for information fusion and dissemination for ground-to-ground service integration. The SWIM ATM added-value services, data access services and technical services defined in Section three are used as a baseline for the definition of the SANDRA Airborne Middleware (SAM) through the utilisation of a set of airborne/ground data domain objects, as described in Section four. Finally, analysis of the service improvement methods and the technology options for future ATM system realisation is addressed at the end of this chapter.

2. SOA in aeronautical communication

SOA is an emerging middleware approach for linking various legacy systems into a standard platform to achieve a highly interoperable and collaborative communication infrastructure. It permits the separation of legacy system service interfaces from the underlying implementation, thus to reduce technology-dependent attributes of the system. SOA promotes service reusability and interoperability through a set of standardised data schemas used in the discovery and self-description of each course-grained, composed and loosely-coupled service.

The SOA capabilities are seen as an enabler for the realisation of the EUROCONTROL SESAR concept (EUROCONTROL, 2008), which explicitly states the focus on the global ATM interoperability with regard to the semantics of the data exchanged, the protocols and the overall quality of dialogue in terms of Communication, Navigation and Surveillance (CNS). According to the ICAO 2010 operational opportunity report (International Civil Aviation Organisation [ICAO], 2010), the European ATM Master Plan defines the "path" towards achieving performance goals by adopting the service-oriented architecture as agreed at the European Union ministerial level. The main targets are to:

- Enable a 10% reduction in CO_2 emissions per flight
- Reduce ATM costs by 50%
- Enable a 3 times increase in capacity
- Improve safety by a factor of 10

Supported by state-of-the-art and innovative technologies designed to eliminate fragmentation in the future European ATM system, SOA-based middleware reflects the operational, technological and regulatory requirements of the future ATM infrastructure while serving for the improvement of system efficiency and interoperability.

2.1 Middleware and service-oriented architecture
2.1.1 Definition of middleware

The term middleware was first introduced in 1968 and had only gained its popularity until it was formally used as an integration platform to connect different systems and applications since the 1980s. The role of middleware varies in different domains. In the scope of enterprise applications integration, middleware is called plumbing because it connects two applications and passes data in between (Simon, 2003). For purpose such as data integration of heterogeneous networks across different geographical locations, especially in the Air Traffic Management context, middleware is a distributed software layer that sits above the operating system and below the application layer and abstract the heterogeneity of the underlying environment. Middleware provides an integrated distributed environment whose objective is to simplify the task of programming and managing distributed applications (Campbell et al., 1999).

The common types of middleware are Message-Oriented Middleware (MOM), adaptive and reflective middleware and transaction middleware. Middleware can be grouped according to different technologies, such as grid middleware, peer-to-peer middleware and real-time middleware concerning the Quality of Service, security, performance, Model-Driven Architecture, Service-Oriented Architecture and more (Mahmoud, 2004).

In aviation, the shift from proprietary air traffic control systems into a standardised and interoperable platform embracing the middleware approach facilitates the communication and integration of a wide range of ground-based and air-to-ground system applications operating across the networks. The middleware acts as an intermediary enabling direct communication with the legacy technology interfaces, to minimise system dependency.

2.1.2 Definition of service-oriented architecture

As an embodiment of the middleware concept, SOA is a paradigm for the integration of various applications running on heterogeneous platforms using common standards. It is designed to consolidate the system capabilities for the organisation and utilisation of data distribution managed by different ownership domains. The term "service" can be defined as a single or multiple operational functions offered by a software system for the fulfilment of business objectives, for example, flight plan filing, aircraft tracking, controller-pilot communication and passenger logistics management. The specification of services may be modified as the business objectives and operations change.

There are eight design principles that affect the design of services and SOA-based system integration (Erl, 2009):

- **Standard Service Contract** – Services within the same service inventory should have the same contract design standards.
- **Service Loose Coupling** – Service contracts ensure the service consumers are decoupled from their surrounding environment.
- **Service Abstraction** – Information contained in the service contracts are limited to what is published.
- **Service Reusability** – Services contain and express agnostic logic and can be reused as enterprise resources.
- **Service Autonomy** – Services exercise a high level of control over their underlying run-time execution environment.
- **Service Statelessness** – Minimised resource consumption by deferring the management of state information when necessary.

- **Service Discoverability** – Services described with metadata can be effectively discovered and interpreted.
- **Service Composability** – Services are effective composition participants.

The SOA concept as a recommendation for system integration is an emerging approach in the ATM development programmes in Europe and North America. It offers a uniform platform, which supports the registration, discovery and administration of individual business process with use capabilities to produce desired effects consistent with measurable preconditions and expectations in a short timeframe.

Rooted in the Business Process Management (BPM) notion, SOA is a holistic mechanism for the alignment and harmonisation of an enterprise and its IT development as:

- SOA encompasses the tools and methodologies for capturing business design, and uses that design information to help improve the business.
- SOA covers the programming model, tools, and techniques for implementing the business design in information systems.
- SOA contains the middleware infrastructure for hosting that implementation.
- SOA encompasses the management of that implementation to ensure availability to the business and efficient use of resources in the execution of that implementation.
- SOA encompasses the establishment of who has authority and the processes that are used to control changes in the business design and its implementation.

The SOA principles and standards highlight the significance of loosely coupled and reusable services in the software architecture perspective. Services are independent and are accessed via standardised interfaces as a frontend of the massive network resources. The advantage lies at the transparent communication SOA offers to end-systems (technology-agnostic), and hence, to effectively demonstrate the application of data-centric information sharing. The SWIM infrastructure provides the basis for information exchange between systems based on the principles of SOA.

2.2 SOA implementation technologies

The definition of SOA emphasises that the concept of service-orientation is a paradigm solely. SOA is remarkable for its flexibility allowing many types of system interactions to be performed based on a series of pre-defined architectural patterns. From the functional point of view, classification of business process and service interaction modelling are two dominant motives at system planning and design stage. Afterwards, the realisation of software services supporting these interactions requires the state-of-the-art technologies to be defined. Technology-independence is one of the most important criterions in terms of technology evaluation.

The paragraphs below provide a general overview on common technologies for the implementation of a service-oriented architecture of which are used in aeronautical communication.

2.2.1 BPM, BPMN and BPEL

In the past 30 years, the growing concept of **Business Process Management (BPM)** has shifted from the use of static business process flowcharts in unchanging organisations to dynamic corporate processes which can be accessed by collaborating partners in a more flexible and efficient way. A business process can be summarised as a collection of structured activity to produce a specific business service or product. BPM introduces

sophisticated software and best practices targeting the modelling, simulation, automation and management of cooperative operations with dynamic business priorities.

Business Process Modelling Notation (BPMN) is a visual language with graphical notations for the modelling of business processes. It presents the business activities, tasks and their relations in a business process diagram (Juric et al, 2008). BPMN can be used in collaborative processes and internal business processes. The BPMN models in future aeronautical communication should appear as a mixture of both intra-business and inter-business flows reflecting the concept of Collaborative Decision Making (CDM).

To implement the modelled business interactions, **Business Process Execution Language (BPEL)** enables the service orchestration for composed service and business processes reinforcing service reusability and loose coupling. BPEL conducts the orchestration of services. It is mapped from the BPMN diagram for service execution, and thus allowing integrated monitoring functions to be applied. The predominant standard for BPEL is the Business Process Execution Language for Web Services (BPEL4WS v1.1) in 2003 (Andrews et al, 2003), defined in the human-readable Extensive Mark-up Language.

Based on the BPM-related concepts, the SESAR SWIM Program has addressed the need to derive a common view on the ATM business processes for accessing SWIM ATM Data and ATM functionalities. The development of a formal business process model has been recommended and it is essential to follow standard IT industry practices through the use of enterprise architecture modelling techniques.

2.2.2 Web services

The Web Service infrastructure is one of the most common approaches for the realisation of SOA. It is recognised as a predominant technology framework in the avionic industry for the realisation of the SWIM infrastructure. The World Wide Web Consortium (W3C) defines Web services as a standard software system for interoperating different software applications running on a variety of platforms and/or frameworks (W3C, 2004). A Web service supports interoperable machine-to-machine interaction over a network based on the Web service stack illustrated in Fig. 1.

Discovery (UDDI)

Description (WSDL)

Standard Messaging Structure (SOAP)

Encoding (XML)

Transport (HTTP)
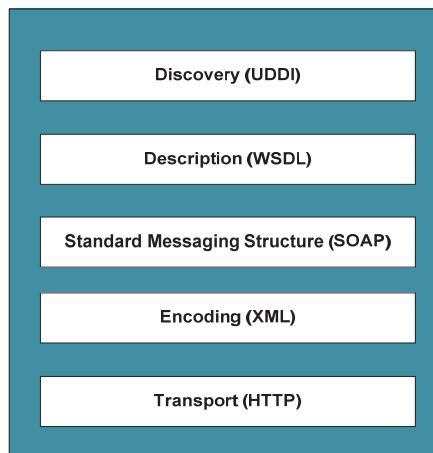
Fig. 1. Web Service Stack.

- **HTTP –** Hypertext Transfer Protocol (HTTP) located in the lowest level of the Web service stack supports the distribution of information across networks. Web applications usually use HTTP on port 80 and HTTPS on port 443 for security purposes. Web service uses SOAP as a messaging protocol over HTTP.
- **XML –** Extensible Mark-up Language (XML) data formatted with XML tags. Data exchanged between services could be described in XML Schema Definition (XSD) conforming to the SOAP as a messaging standard.
- **SOAP and transport protocols –** Simple Object Access Protocol (SOAP) as a standard messaging protocol, and other transport protocols such as HTTP, FTP, SMTP and JMS are common protocols for web service communication.
- **WSDL –** Web Service Description Language (WSDL) is an XML formatted file used to abstractly describe the network services corresponding to the endpoints. WSDL defines the service details (contracts) between service consumers and providers. Messages, port types for specific operations, bindings, services containing SOAP addresses are key elements of a WSDL file.

Web service stands out for its standardised languages and specifications. Nevertheless, the verbose structure and the long processing time of XML schema are two major constraints. Therefore, it is essential to define suitable mechanisms to improve the efficiency of XML applied in aeronautical communication where the performance level should be optimised in order to ensure the Quality of Service. As addressed in Section 4, the matching of Abstract Syntax Notation One (ASN.1) and XML and compression mechanisms can be considered as potential solutions.

2.2.3 Message-oriented middleware

Message-Oriented Middleware (MOM) is a software term that connects multiple systems in a network for data distribution, typically in a service-oriented architecture. It is inspired from the conventional client/server architecture nevertheless, with advanced features allowing both synchronous and asynchronous communication. Such feature is beneficial in providing a messaging platform accessed by global ATM systems.

The introduction of MOM enhances the level of quality attributes such as performance, scalability, flexibility and interoperability. MOM provides Java Message Service (JMS) as the standard Application Programming Interface (API) to perform the message broker function and to support client application.
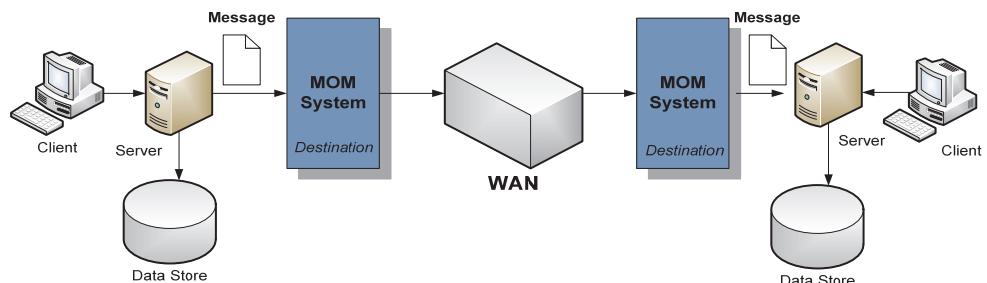
Fig. 2. MOM system architecture.

As indicated in Fig. 2, a MOM system can deliver message across networks via a centralised message server or it can distribute routing and delivery functions to each client machine. The client can continue requesting information from the data store or performing other operations while delivering messages to the JMS API.

2.2.4 Enterprise service bus

An Enterprise Service Bus (ESB) is a common implementation solution that allows services to be used in a productive system. An ESB provides coarse-grained interfaces with the purpose of sharing data asynchronously between applications. Such integration architecture pulls together applications and discrete integration components to create assemblies of services to form composite business processes, which in turn automate business functions in a real-time enterprise. The rise of multiple ESBs is a result of iterative SOA implementation approaches.

An ESB provides (but not limited to) the following functions:

- Messaging functions, such as transformation, delivery and routing
- Service registry and metadata management for the storage and discovery of services
- Adapter functions supporting various communication protocols
- Support to allow service composition/orchestration in business processes through WS-BPEL.
- Security management in order to provide authorization, authentication, and the creation of the policies
- Management monitoring and configuration of the management, components life cycle, logging and auditing

2.2.5 Data distribution service

Data Distribution Service (DDS) is a newly adopted middleware specification for distributed real-time applications introduced by Real-Time Innovations (RTI) in 2003. It is a standard implementation of Object Management Group (OMG) and is used in many time-critical and data-critical applications such as industrial automation, robotics, air traffic control and monitoring and transaction processing.

DDS is aimed at a diverse community of users requiring data-centric publish/subscribe with high flexibility, performance, portability and configurable data distribution management using the topic channels. Such publish/subscribe based communication model is used for sending and receiving data, events, and commands among the service nodes managed by different publishers (containing any number of DataWriters) and subscribers (containing any number of DataReaders) connected to the Global Data Space for resources sharing.

The development trend of DDS, e.g. the amalgamation with Web Services standards is driving DDS to a maturing SOA solution for future aeronautical service integration.

2.2.6 Common object request broker architecture

Common Object Request Broker Architecture, namely CORBA, is a specification defined by the OMG for system integration of aeronautical legacies. However, as the Web Service based solutions are gradually gaining recognition, CORBA is shifting to a legacy category.

2.3 SOA and air traffic management

The investigation and analysis of SESAR SWIM-SUIT (System-Wide Information Management SUpported by Innovative Technologies) and the FAA SWIM prototypes

reveal substantive findings on service integration of ground-to-ground communication in a service-oriented architecture (EUROCONTROL, 2008; FAA, 2010). Implementation of the functional framework is carried out using Web Services, JMS, DDS and ESB to support one-to-one, one-to-many and event-based communication via request/reply and publish/subscribe message exchange patterns. It is envisaged that the SWIM-based SANDRA Airborne Middleware and future air traffic management infrastructure will also consider SOA as a baseline for seamless aeronautical communication in the next decades.

3. The SWIM SOA approach

3.1 SWIM overview

System Wide Information Management, namely SWIM, is an information sharing concept leading to the development of a number of technology programmes conducted by both the FAA, as the Next Generation Air Transportation System (NextGen), and EUROCONTROL to facilitate the information sharing and global situation awareness in the future aeronautical context. In the past, the state-of-the-art to connect different ATM systems required a fixed connection and application-level data interfaces to be set up individually between each system. SWIM is essentially introduced to reduce the high degree of system dependence by providing a Network Centric environment to enhance the flexibility of aeronautical system integration.

ICAO Global ATM Operational Concept definition of SWIM (SWIM-SUIT, 2008):

"System Wide Information management aims at integrating the ATM network in the information sense, not just in the systems sense. The fundamental change of paradigm forms the basis for the migration from the one-to-one message excha is geographically dispersed sources collaboratively updating the same piece of infornge concept of the past to the many-to-many information distribution model of the future, thatmation, with many geographically dispersed destinations needing to maintain situational awareness with regard to changes in that piece of information. Successfully managing the quality, integrity and accessibility of this complex, growing web of distributed, fast changing, shared ATM information, called the virtual information pool, can be considered as the main operational enabler for the operational concept."

FAA description of SWIM (SWIM-SUIT, 2008):

"To streamline the evolution and modernization process, SWIM concept is to support loosely coupled, many-to-many data exchange interfaces. When implemented, SWIM will allow information producers and consumers to exchange data in a secure, robust, standards-based, loosely coupled environment." […] "Exploitation of advancing technology that moves from an application centric to a data-centric paradigm – that is, providing users the ability to access applications and service through web services – an information environment comprised of interoperable computing and communications components."

The EUROCONTROL SESAR definition of SWIM (SWIM-SUIT, 2008):

"SWIM represents added value also in terms of facilitating general accessibility. Focus shifts from the producer of information to information itself and generalised access to information (as opposed of pre-packaged sets as is the case today) enables users to create their own applications which best suit their mission needs. In the ATM network, almost every participant is a producer as well as a consumer of information. It is not desirable to decide in advance who will need what information, from whom and when. The key issue is to decouple producer of information from the possible consumer in such a way that the number and nature of the consumers can evolve through time. On the contrary for what concerns the producers of information it is of the utmost importance to agree on the level of interoperability required with other ATM stakeholders that may have to contribute to the elaboration

of the consistent and consolidated view of the reference data. For that purpose, the SWIM participants have to share:

- *A reference Data and Services model,*
- *A set of agreed cooperation patterns (Rules, Roles and Responsibilities),*
- *A set of technical services necessary to support system interactions,*
- *An access to the SWIM physical network.*

In short, SWIM provides the mechanisms which support the partners in managing the Rules, Roles and Responsibilities (the 3Rs) of information sharing. This determines which kind of information is shared by whom, with whom, where, when, why, how, how much, how often, at which quality level, in what form, for which purpose, at which cost, under which liability, under which circumstances, security level of air traffic management. The 3Rs must also be properly addressed both in terms of institutional and Information Communication Technology (ICT) aspects.

3.2 SWIM in Europe

Since 1997, EUROCONTROL's Experimental Centre has been participating in the SWIM-SUIT Project, an initiative laying some of the foundations of SWIM. In 2008, the EUROCONTROL launched SWIM-SUIT as an underlying work package of SESAR with the aim to facilitate aeronautical information sharing with regarding to flight data, airport operational status, weather information and special use of airspace and restrictions. The information sharing targets a number of operators working with the ATM systems in aspects such as airline operation control, administration, air traffic services and passenger and logistics management.

Building on top of the SWIM concept, the SWIM-SUIT Project was designed to allow expedite and secure access and conveyance of vital information supported by the process of collaborative decision making with the adoption of the state-of-the art technologies, for achieving efficient and effective cross-domain operations.

3.2.1 Scope and timeline

The SESAR Concept of operations for the long-term time frame calls for an overall European ATM system (EATMS) that is fully interconnected via a SWIM network. As illustrated in Fig. 3, the connected systems include:

- Pan-European systems for managing Europe/network-wide information services;
- Civilian and Military ATC systems;
- Airspace users systems (Military, Scheduled and on-demand civilian operators);
- Airports;
- Aircraft.

Such a system (or rather a "system of systems") requires a move from point-to-point message exchange to the sharing of information within a common virtual information pool. The SWIM architecture will permit actors to focus upon the information itself, rather than the systems that produce / manage the information.

The SESAR SWIM environment, given its wide reaching scope, will require a progressive, iterative and constant implementation programme. The following diagram outlines the current SESAR view on the various developments streams (hence, the deployment begins progressively following the end of a development stream).
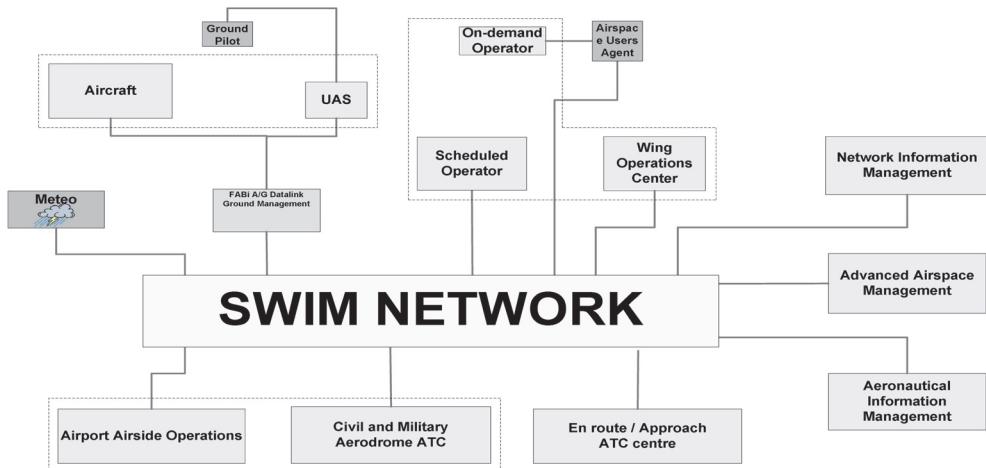
Fig. 3. SESAR SWIM architecture.
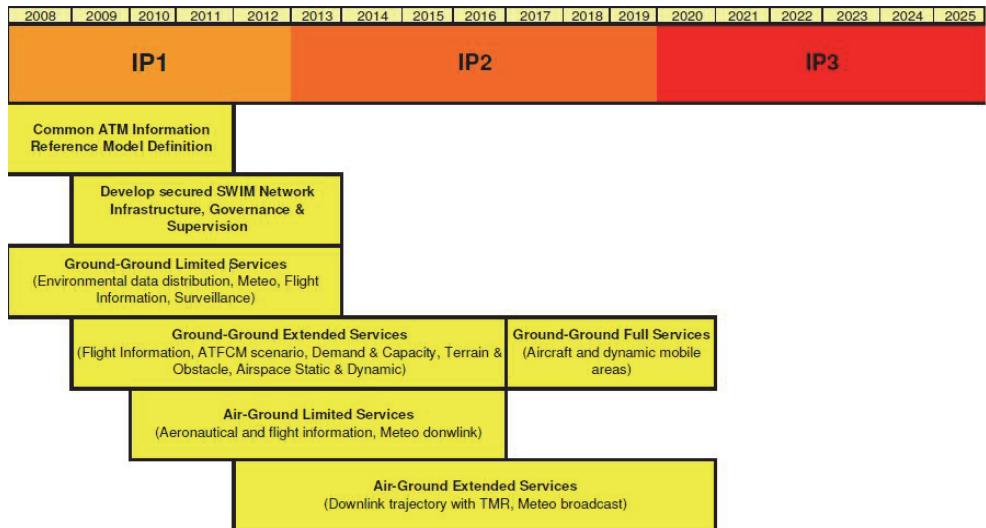
Fig. 4. SESAR SWIM implementation timeline.

3.2.2 Data sharing and services

The SESAR SWIM environment expects at least the following data to be shared across the SESAR-defined time frames, known as the short-term/medium planning, long-term planning, and the execution and post-execution phase:

- **Flight Data** - Flight Structure, Flight Script, Taxi Plan, Trajectory, What-If Flight and Context, and departure and arrival related data represented in the Flight Object Model (The European Organisation for Civil Aviation Equipment [EUROCAE], 2006)
- **Surveillance Data** - System Track, Sensor Descriptions, Aircraft Track, Traffic Advisory and ASAS Alert

- **Aeronautical Data** - Aerodrome Data, Heliport Data, Airspace Data, Navigation aids Data, Terrain and Obstacles Data and Aircraft Data with details described in (Aeronautical Information Exchange Model [AIXM], 2011)
- **Meteo Data** - Aerodrome Weather, Area Weather and Met Hazard with details described in (Weather Information Exchange Model [WXXM], 2011.)
- **Capacity and Demand Data** - Configuration Plan, Traffic and Airspace Demand, Traffic Load and Demand Capacity Balancing Measures
- **ATFCM Scenario Data** - Demand Capacity Balancing (DCB) Scenario, Flow Measures and DCB Scenario Catalogue

3.2.3 System architecture

Fig. 5 illustrates a layered conceptual view of the SWIM architecture:

- **SWIM network –** the physical pan-European network along with the essential technical IP building blocks (e.g. transport protocols, firewalls).
- **SWIM Technical Services –** the core technical services that SWIM will provide to all connected systems. These services are built, as far as possible, upon standard IT middleware technologies.
- **SWIM ATM Data Access Services –** this layer embodies the "SWIM virtual information pool". It provides access via defined services to the standardised SWIM ATM Data model. The data access services are typically be categorised into different domains (e.g. FlightDataAccess, MeteoDataAccess).
- **SWIM ATM Added-value services –** this layer contains services that provide access (perhaps distantly) to added-value ATM functionality beyond that of the "virtual information pool". Typically, this layer could contain CDM type applications/services.
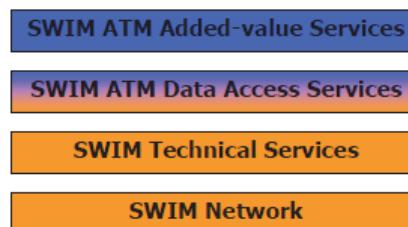
Fig. 5. SWIM layered architecture.

From a domain viewpoint, the layers can be divided into two main groups:

- **SWIM Infrastructure –** SWIM Network, SWIM Technical Services, and SWIM ATM Data Access Services (partial), which represent a common SWIM IT infrastructure, consisting of both turn-key solutions and toolkit/frameworks, upon which is built the SWIM ATM functionality.
- **SWIM ATM functionality –** SWIM ATM Added-value Services, SWIM ATM Data Access Services (partial), representing the domain specific functionality.

The SWIM infrastructure is spread out amongst the (legacy) ATM systems that form a part of the overall European ATM system (the "system of systems"). Each ATM system, typically composed of a number of major subsystems, implementing domain specific functionality will now include a "SWIM / IOP Management subsystem" which:

- Contains the (or parts of) SWIM Infrastructure functionality shown above;
- Connects the legacy system functionality to the SWIM environment;
- Translates standard SWIM ATM data structures into the appropriate legacy system data formats.

Services in SWIM are defined in a domain-specific manner providing a wide range of standard functional interfaces to support the communication and collaboration between participants connected to the SWIM network. Data collected from both the SWIM-enabled applications and ATM legacy systems, via the SWIM external adapter, is transmitted through the SWIM Ground/Ground gateways, namely the SWIM-BOXes, for the facilitation of data sharing.
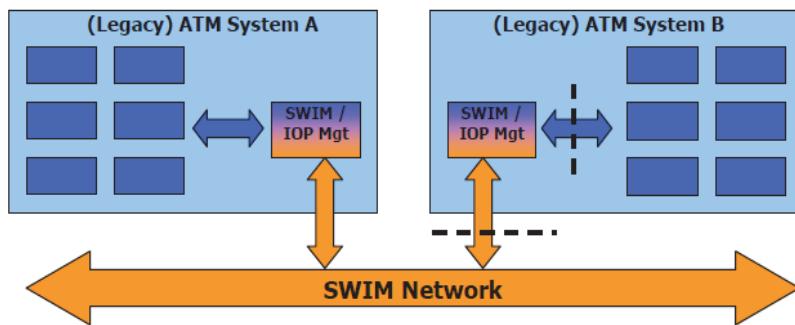
Fig. 6. SWIM/ATM system viewpoint.

3.3 SWIM in the U.S.

The SWIM Program in the United States was originated in 1997 as the EUROCONTROL initially presented the SWIM concept to FAA. The concept was under development ever since until the ICAO Global ATM Operational Concept adopted the SWIM concept in 2005. SWIM is now a part of development project in the United States NextGen framework for the development and integration of the National Air Space (NAS) systems for greater sharing of ATM system information on airport operational status, weather information, flight data, status of special use airspace, and NAS restrictions.

3.3.1 Scope and timeline

The FAA has established a notion called "Core Services" as a consistent capability existing at each node to provide a uniform mechanism for communicating among nodes. Fig. 7 illustrates the FAA view of how Core Services fit into the overall SWIM architecture.

The following system cores are included:

- En Route Automation Modernization(ERAM)
- Weather Message Switching Centre Replacement (WMSCR)
- Traffic Flow Management System (TFMS)
- FAA Telecommunications Infrastructure (FTI)
- Special Use Airspace Management System (SAMS)
- Central Processor (CP)
- National Airspace System (NAS)
- Electronic Flight Strip Terminal System (EFSTS)

- Airport Surface Detection Equipment –Model X (ASDE-X)
- Flight Data Input Output (FDIO)
- Terminal Data Link System (TDLS)
- Runway Visual Range (RVR)
- Air Route Traffic Control Centre (ARTCC)
- William J Hughes Technical Centre (WJHTC)
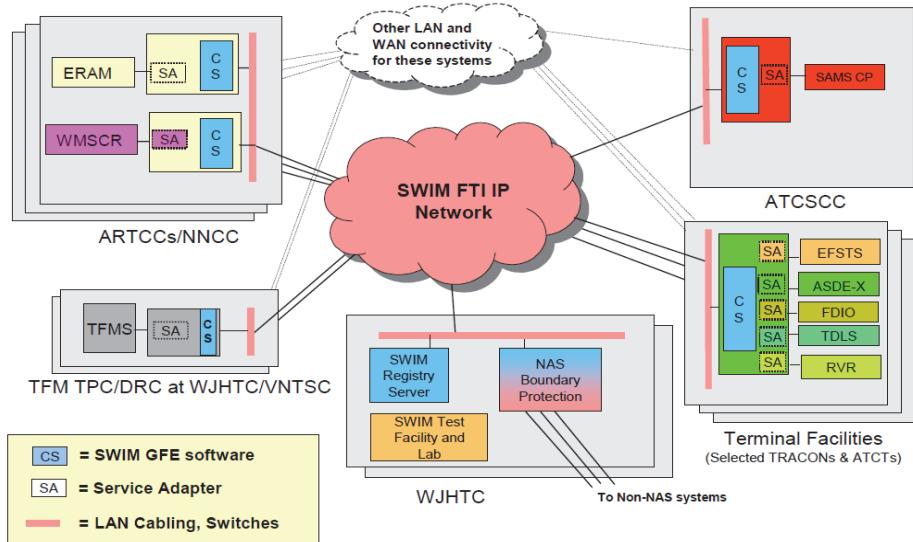
Fig. 7. FAA SWIM architecture with core services.

Fig. 7 raises the question of what specific capabilities are comprised in these Core Services. The FAA has proposed the core capabilities in Fig. 8 below.
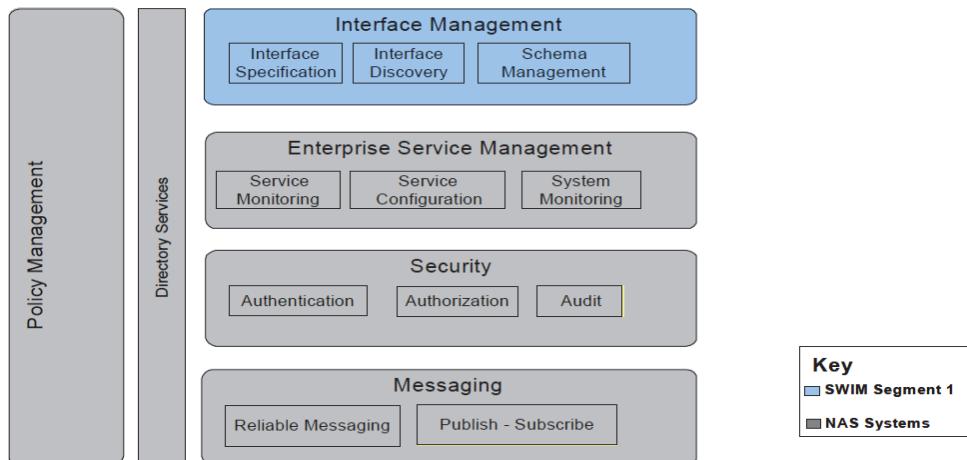
Fig. 8. SWIM Segment 1 Core Capabilities.

The FAA organizes SWIM's initial Core Services into four groups (SESAR shares the same Core Services grouping as FAA):
- Interface Management
- Messaging
- Security
- Enterprise Service Management

FAA SWIM will be deployed in segments (stages), with the first segment planned for the 2008-2012 timeframe though the NextGen has a long planning horizon (20+ years).

3.3.2 Data sharing and services

The Segment 1 business services are defined in the SWIM Final Program Requirements. It identifies and describes the services in the following categories for example:

Flight and Flow Management
- Flight Data Publication
- Terminal Data Publication
- Flow Data Publication
- Runway Visual Range (RVR) Data Publication

Aeronautical Information Management
- Special Use Area (SUA) Data Publication
- Corridor Integrated Weather System (CIWS) Data Publication
- Integrated Terminal Weather Service (ITWS) Data Publication

3.3.3 System architecture

In response to the FAA SWIM program using SOA, Government Electronics & Information Technology Association (GEIA) which is a trade association that includes many industry partners who support the FAA, provides the industry solution, shown in following figure, both adheres to the overall SOA and to the FAA SWIM vision.
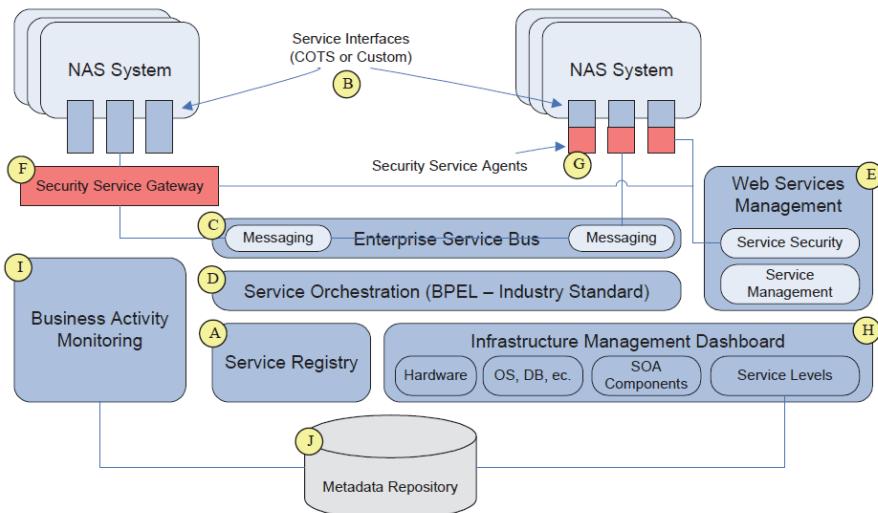
Fig. 9. GEIA SOA Architecture for SWIM (GEIA, 2008).

The GEIA architecture above supports the FAA Service groupings via the following subsystems, according to the "SOA Best Practices –Industry Input" (GEIA, 2008).

- Interface Management (A, B)
- Messaging (C, D)
- Security (E, F, G)
- Enterprise Service Management (H, I, J)

4. SANDRA – extending SWIM onboard

4.1 SANDRA overview

Building on the SESAR SWIM concept for information fusion and dissemination for ground-to-ground service integration, the EU FP7 project SANDRA (Seamless Aeronautical Networking through integration of Data-Links, Radios and Antennas) extends the ideology of SWIM to cover air-to-ground information exchange, service composition and integration to provide a complete and coherent set of communication services for future global Air Traffic Management in the 2020 timeframe. To ensure the integration of different service domains onboard legacy applications with very diverse requirements, the SANDRA communication system will represent a key enabler for the global provision of distributed services for Collaborative Decision Making based on the SWIM concept, and for meeting the high market demand for broadband passenger and enhanced cabin communication services. As a case study, the paragraphs below concentrate on the SANDRA Airborne Middleware design and how such architecture can be realised using the various SOA technologies.

Focusing on communications related aspects, the following high-level requirements are identified to allow future systems to be compatible with the expected air-traffic growth:

- Pilots situation awareness shall be improved
- Capacity at airports, as today's main limiting structural factors, shall be increased
- ATS shall be primarily based on highly reliable data communication
- AOC data traffic shall strongly increase for efficient airline operations
- Passengers and cabin communications systems shall be further developed
- Safety critical applications shall need diverse means to reach ground for global availability and higher reliability
- A simplification of on-board network architecture shall need convergence of protocols and interfaces

In order to satisfy the objectives (middleware aspect only), a possible airborne middleware architecture in SANDRA is defined aiming at the interoperation with the ground systems utilising SOA. To define the airborne middleware architecture, analysis of air/ground (A/G) information exchange is carried out focusing on the definition of the A/G data domain services and functional interfaces based on the SWIM information infrastructure. A combination of mechanisms to improve the A/G data exchange is proposed in dealing with the limited bandwidth available for over the SANDRA Data Link. The A/G integration architecture is defined containing a set of core infrastructural subsystems.

4.2 Analysis of air/ground information exchange

The SESAR SWIM environment expects at least the following data to be shared:

- Flight Data
- Surveillance Data
- Aeronautical Data

- Meteo Data
- Capacity and Demand Data
- ATFCM Scenario Data

4.2.1 SWIM ATM Information model

The ATM information to be exchanged via SWIM Network needs to be modelled explicitly, to allow a precise and concrete definition to be agreed. Previous work has already defined data models and required services within specific domains (e.g. Aeronautical Information and Flight Information), but this is not the case for all types of information. The services in support to SWIM are organized around a number of central themes called profiles.

An overall ATM Information Reference Model is required to define the semantics of all the ATM information to be exchanged. This model will form the master definition, subsets of which would be used in lower level models, supporting interoperability for each of the data-sharing domains.
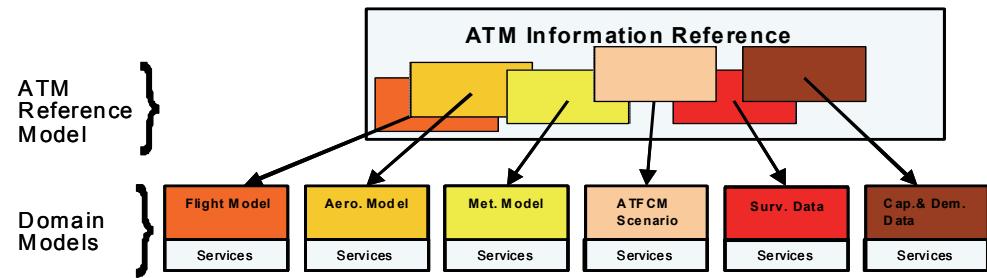
Fig. 10. SESAR ATM Reference Information Model.

The ATM Information reference would be a key asset in the ATM System design and would sit above a set of domain-specific, platform independent models which may overlap with each other, without being incompatible. The overall reference model and existing models such as OATA, FOIPS, OMEGA, AIXM etc., will need to be reconciled.

From these models, the specific SWIM exchange formats (i.e. on the wire exchange formats) can then be defined. SWIM assumes that the data handed over to SWIM Node services is already in the canonical ATM information model. There will be no mappings from/to the canonical data model in the data domain components. If such mappings are required, they have to be implemented in the SWIM Applications adapters. The data domain components perform data conversions for encryption/decryption and data filtering on the canonical data.

4.2.2 SWIM ATM information services

The ATM information services are all services that can be called on the northbound external interface of a SWIM Node. These are application services and data access services.

Fig. 11 shows a typical scenario: A simple request/reply service invocation results in a data change (on ATM System B) that is followed by a publish/subscribe exchange to distribute the changed data. In this case the ATM information service call is initially triggered by the SWIM Application Adapter on System A. System B is identified as the master of the data and the service call is forwarded to System B where the operation is performed. If the operation results in a data change, this is distributed to all relevant subscribers.

Only the initial service call originates outside of SWIM, all the other information flows are triggered by the SWIM infrastructure. This requires some sort of service orchestration inside of SWIM.
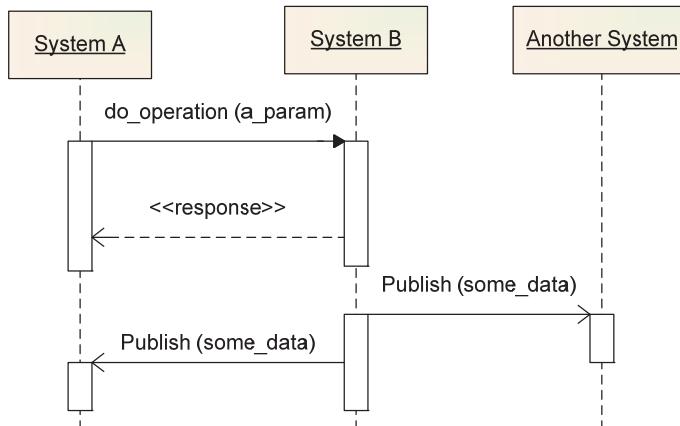
Fig. 11. Interaction Scenario.

4.2.3 SANDRA data domain concept

In order to respect the SWIM Data Domain Concept, described in SESAR as "Profiles", the SANDRA Airborne Middleware (SAM) provides a more generic approach to allow different kind of data sharing/services over SANDRA data links without any kind of restrictions, but just working on XML Optimised Representations of Shared Data.

Fig. 12 below clarifies the Data Domain Object (DDO) concept: as the Flight Object on SWIM Ground Side is described as composed by different Flight Object Clusters with an XML Representation, SANDRA DDO provides a more generic view of the same thing, in particular it is composed by different Data Object Clusters and, with this assertion, we can describe any kind of Data Domain (Meteo, Aeronautical, Flight, etc) using this kind of representation for specific entities.

For example, SANDRA supports the EUROCAE ED133 Flight Object services (EUROCAE, 2009) through the DataDomainObject.

4.3 Improving the air/ground data exchange services

SANDRA will provide a generic representation for the shared data over SAM, in particular, analysing the last image it is possible to discover two kinds of representations for the same information:

- **DataDomainObject (DDOject)**: it contains "n" DDObjectClusters of each with an Object Identifier and multiple Object Release Identifiers described in XML
- **DataObjectSummary**: it contains just one data object cluster the Distribution List cluster that contains information about the participants over the DDObject shared information. A release identifier, contained also in the DDObject, gives information about the actual releases of the single clusters contained into the DDObject.

These two representations are important in order to reduce the traffic between the involved stakeholders: the first one is the full data representation that each involved stakeholder has
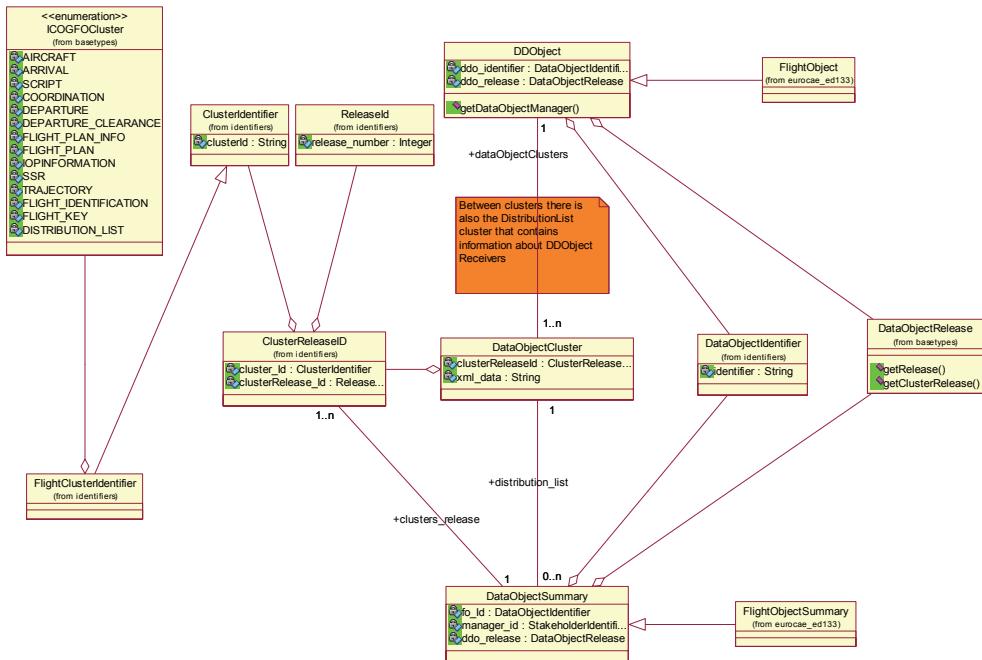
Fig. 12. SAM Data Domain Object.

to save locally; the usually shared information that informs which is the last "present" information is given by the DDObjectSummary.

Once the new AIRBORNE requests a subscription, it receives all the summaries of DDObject of interests (based on particular AIRBORNE StakeholderIdentifier). The SWIM Ground infrastructure may need to be extended in order to identify the objects exchanged to support SWIM A/G communication. For example, it may be useful to link the specific DDObjects (Meteo information, Flights over the trajectory, etc) to a Trajectory for distribution of the DDObjectSummaries of interest to subscribed aircrafts only.

SANDRA targets to reduce the traffic on the Air/Ground data link by designing a more schematic yet lightweight communication protocol. However, a publish operation for distributing the whole clusters information in XML format incurs a large overhead and high bandwidth usage. In order to reduce the amount of shared information, two important operations are applied to the XML format:

- Compression: A number of compression techniques can be used to reduce "drastically" the XML shared information size.
- Data Manipulation: the XML format is extremely schematic and there are techniques that can allow simple updates over an original data.

In considering the above, the requirements for data exchange optimisations are identified as follows:

- A common data representation to optimise the shared data is required
- An XML-based shared data content is required
- A complete knowledge of Shared Information by the ground system is required in order to share essential information starting from an airborne stored one.

To respect these few but extremely important requirements, the DDObject with its inner XML clusters are introduced. Concerning the Ground side DDObjects knowledge, a "DataRepository" component to store all the DDObject Cluster Releases of each DDObject is defined. The "Delta" concept for the differentiation and integration of XML data is applied using the information stored in the "DataRepository".

Based on the SESAR SWIM study, the "publish" and the "read" operations are the two most critical operations to integrate compression and data manipulation concepts into SAM:

- **Applying "Delta" Concept**: Once the DDObject Manager requires an update on an already shared DDObject, it requests a publish operation of the DDObjectSummary that contains the DDObject Identifier, DDObject Release and Distribution Cluster for such information to be received by airborne applications that have already been subscribed to and registered on the DDObject Domain, also given that the DDObject is created and updated. Under this premise, the airborne application knows the "stored" DDObject Release and the new Release Identifier that was just published. Knowing this information the application can request a "read" operation signalling the "stored" release to enable the ground side to compare the release values of the newly published DDObject and the one that is stored by the airborne application. Once the SWIM Ground locates the clusters identifier to update using a DDObject repository that contains all the DDObject Releases clusters, it can be possible to identify the differences between the XML file stored in the airborne application and that in the newly shared clusters. The difference in the XML contents, XML Diff, can then be sent as a reply message to the read operation signalled by the airborne application.

- **Applying Compression Concept**: Compression can be defined at different layers, in particular at the "message" layer and at the "parameter" layer. In order to design the SANDRA Middleware, Web Service technology and, in particular, the shared messages – SOAPMessages to realise the request/reply pattern are applied. Previous studies show that it is possible to compress SOAP messages before sending out the requests and to decompress them once received at the destination. In addition, with the "Delta" Concept, the SOAP message body will contain only a "read reply" message with just the differences between the "Airborne Stored DDObject Clusters" and "Ground Updated Clusters" as parameters. As a result, the SOAP Body contains only the XML Diff content. Applying XML Compression over XML Diff will further reduce the shared packet size.

The same compression can be however applied to the publish operation since all the DDObjects has the same data structure XML Based. Hence, it is possible to compress the publish topics clusters - the whole clusters for DDObjects and just the Distribution Cluster for the DDObjectSummaries before publishing.

4.4 The SANDRA airborne middleware architecture

The SANDRA Airborne Middleware architecture as shown in Fig. 13 has the scope to provide services to airborne applications and to the ground-based Air Ground Data Link Ground Management System (AGDLGMS) gateway over the SANDRA A/G Network. In order to establish such a connection, a set of external interfaces are defined for use in the ground side to communicate with SWIM Ground Gateway and airborne side to realise the communication with SWIM airborne applications.

A detailed SAM system architecture, as illustrated in Fig. 14, includes the AGDLGMS entity residing on the ground is communicating with the other SWIM G/G Gateways via the SWIM Ground Network; and also the SANDRA Airborne Middleware, via the SANDRA A/G Data Link. The SAM provides two kinds of external interfaces:

- **SAMUtilityService Interface:** this interface has to provide the utility services requests such as Subscription, Un-subscription and the set of connected Airborne status
- **SAMBusinessService Interface:** this interface has to provide the business services requests, in particular creation, update, handover, read, and so on.
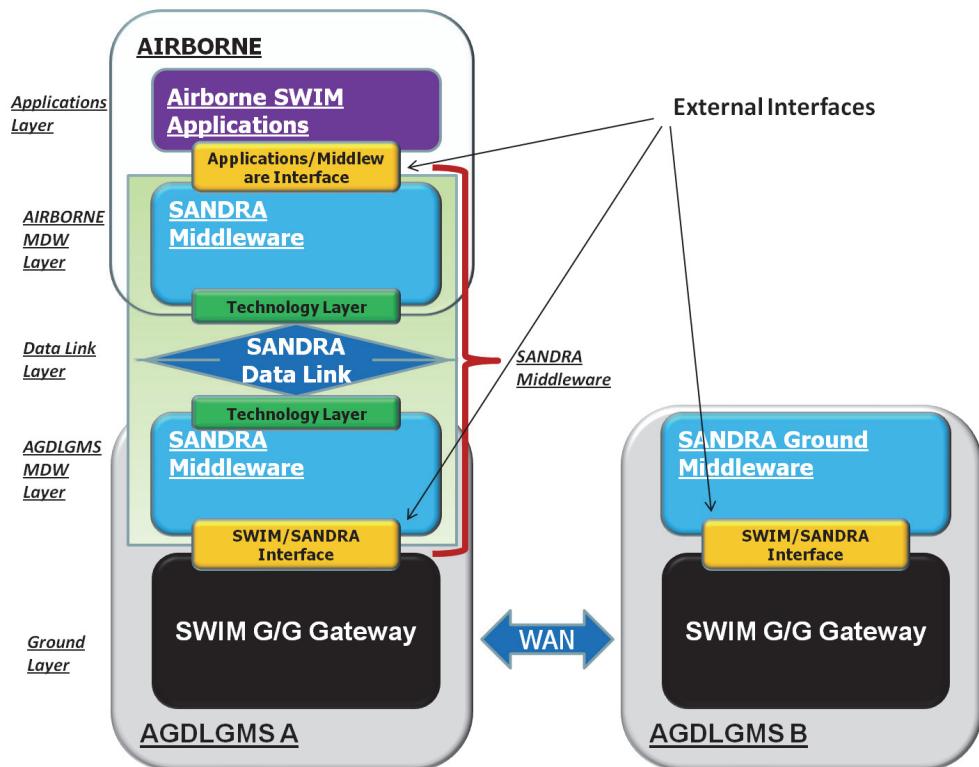
Fig. 13. SAM Architecture Overview.
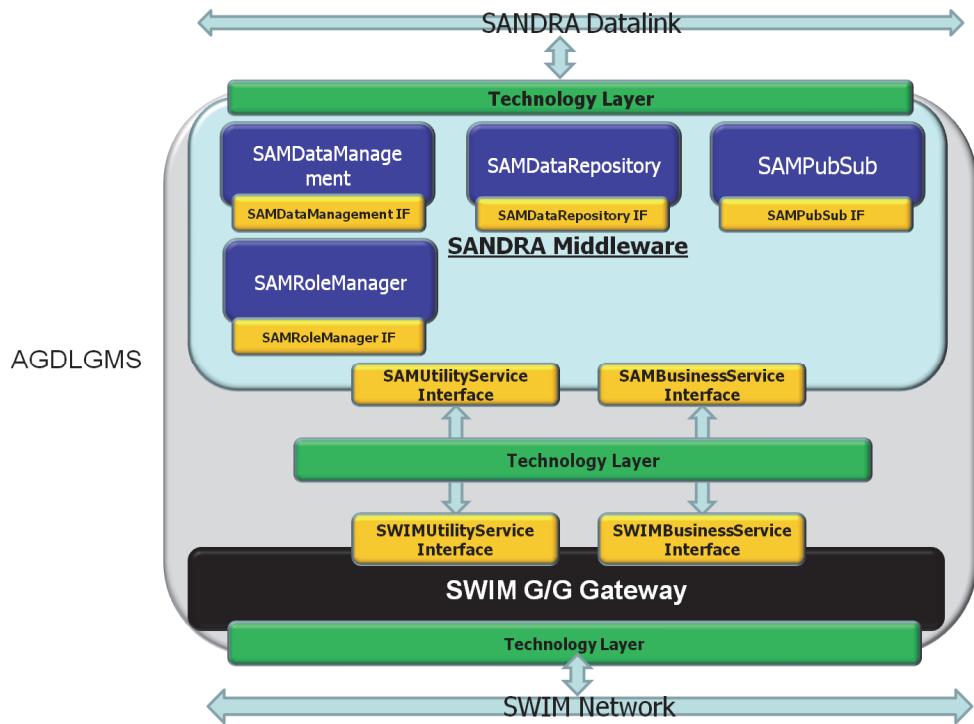
Fig. 14. SAM Architecture Details.

As described in the following paragraphs, both the airborne and ground SANDRA middleware segments should provide the core infrastructural functions to support the aeronautical operations via the defined SAM interfaces. Core services are defined in alignment with the EUROCONTROL and FAA SWIM approach to construct a coherent SWIM infrastructure around the globe.

4.4.1 Interface management

The interface management provides capabilities that enable service providers to publish information about services and service consumers to discover service information. In addition to the generic service registration/publication and discovery, user registration and subscription for notification are also associated supporting functionalities.

The interface management of the SANDRA system should support the manipulation of information from the service registry, shown as follows:

- **Service Registration**: the ability for a user to register a service to the system.
- **Service Lookup**: the ability for a user to look for (discover) a registered service.
- **Service Invocation**: the ability to invoke an identified and possibly distant service.

A service provider registers the service endpoints on a local registry. For routing purposes, the lookup service will be invoked if the service consumer does not know the endpoint of a specific service. However, this also depends on the message exchange pattern being applied. Both Flight Object key selection and content-based search support the service discovery.

4.4.2 Messaging services

The messaging services provided SAM the messaging mechanisms to enable the communication between service providers (publishers) and consumers (subscribers) in a loosely-coupled manner. However, these roles may vary in different scenarios. The decoupling feature of participants not only reduces chances of synchronisation, but also ensures the transparency and autonomy of services. Several mechanisms are introduced here to support a variety of services invocation styles and data exchange protocols:

- **Publish/Subscribe with notification**: allow publishers to check the subscribed users for the actual notification type and send them the subscribed notification.
- **Request/Response**: allows service client to request information from the server through for example, RPC-style client/server communication.

The adoption of Web Services and JMS allows both synchronous and asynchronous messaging for the realisation of various aeronautical operations across all data domains.

4.4.3 Security services

The security services offered by SAM focus on aspects such as user role management, access control, the encryption and decryption of messages as well as service security monitoring. The Lightweight Directory Access Protocol (LDAP) user registry is considered in the design phase to provide security functions.

A common approach is introduced to allow service consumers to gain restricted access only on specific tasks by building a standalone authentication and user management application that can be accessed as a service in a SOA framework. Operational services/ processes that require proof of identity will access this shared user account; thereby the shared function can be repurposed across all tasks. In this scenario, the shared user account service will be reusable for different business processes. This includes the following aspects:

- **Authentication**: this function provides password features for user login.
- **Authorisation**: this function provides user access control.
- **Logging**: this function stores the user actions in a security log and alerts unauthorised operations according to the security requirements (SANDRA, 2010).

4.4.4 Enterprise service management

Enterprise Service Management (ESM) includes the governance and monitoring of services. It provides passive and active management of services at run time while performing overall management operations of systems and applications.

SAM should also provide core functions to collect the Service Level Agreement (SLA) and policy-related metrics to ensure QoS. An SLA refers to a set of performance figures that are indicated in the service contract shared by a service provider and the consumer(s) of that service. Guaranteed operation time, i.e. system uptime, downtime and response time, as well as capacity of the system are the most common figures. The topics covered are:

- **Fault monitoring and reporting**: with this function the administrator can monitor and manage the service faults.
- **Performance monitoring and reporting**: with this function the administrator can monitor the performance level of the services.
- **Enterprise configuration**: with this function the administrator can perform system configuration (e.g. service setup and software configuration) through GUI.
- **SANDRA service management**: with this function the administrator can monitor the quality level of services, according to the QoS requirements (SANDRA, 2010).

4.5 Technology choices

A number of technologies are proposed for the realisation of the SANDRA Airborne Middleware in a service-oriented architecture. Evaluation of the suitable technologies is performed based on the three following steps:

- The propose of candidate technologies
- A list of evaluation criteria taking into account the system requirements and communication patterns of SWIM, SWIM-SUIT and SANDRA Airborne Middleware (SAM)
- Technology selection based on the rating of each technology against the quality attributes to define a technology selection matrix

In summary, it is envisaged that Web Service is the one of the most appropriate technologies in future aeronautical communication considering its standardisation, interoperability and compatibility with the ground SWIM infrastructure. It is also suitable in the airborne context for its maturity and integration features. Web services provide the support for the realisation of a SOAP-based SOA infrastructure incorporating the MOM JMS API for providing pragmatic request/reply and publish/subscribe messaging mechanisms. The Enterprise Service Bus is considered as an alternative approach for system implementation in addition to Web Service and MOM/JMS. DDS outstands for its excellent network performance. Therefore, it is envisioned as an emerging candidate technology for high-efficient and technology-independent data distribution along the refinement of its specification. The SWIM-SUIT project has also concluded with the same prospect and adopted DDS for the SWIM-BOX implementation on ground.

The evaluation of the implementation software indicates that Apache and JBoss frameworks are respectively the most and second-most appropriate technologies. The scores are obtained based on the aggregation of a Web service infrastructure, a messaging backbone (MOM) and an ESB for service composition and system integration. Here, the following sets of technology options are proposed:

- Apache CXF, Apache ActiveMQ, Apache Camel, and Apache ServiceMix or Apache Synapse - Apache CXF and ActiveMQ compose a lightweight system framework for onboard systems. Apache Camel can be used for advanced routing and mediation functions if required. Apache Synapse can be used as a replacement of ServiceMix to achieve a more lightweight system. System bridging is required for achieving full interoperability with SWIM-SUIT.
- JBoss Application Server/JBoss ESB, and HornetQ - direct integration with SWIM-SUIT, thus to achieve full interoperability. HornetQ is claimed to have better performance among all message brokers available.
- XOperator or XMLUnit - used for comparing, summarising and synchronising XML documents for the realisation of the "Delta" data sharing in the SWIM context, i.e. to support operations such as "getDataDiff", "makeDataDiff" and "integrateDataDiff" in the "SAMDataRepository" and "SAMDataManagement".
- FastInfoset and Fast Web Service - used for converting test-based XML into binary-based ASN.1 FastInfoset provided by the JAX-WS interface within Java framework while remaining the capability to allow message transmission using SOAP. No additional tool is required for this approach. Although the support for Fast Web Service has yet become sufficient, the approach is envisioned for SAM development.

5. Conclusion

With an appreciation of envisioning the future aeronautical communication system in a service-oriented architecture, this chapter summarises the ongoing development of avionic service integration and the utilisation prospects of the middleware and SOA concepts in the 2020 timeframe and beyond. SWIM, as an architectural blueprint, is envisaged as a baseline solution in dealing with the growth of airline passenger and the increasing complexity of aeronautical operations. The state-of-the-art SWIM system architectures proposed by world-leading aviation organisations, e.g. EUROCONTROL and FAA allow Air Traffic Management stakeholders around the globe to obtain coherent and persistent knowledge of relevant system data regarding the flights in operation.

To continue with the ground-based SWIM programmes, this chapter proposes a possible SANDRA Airborne Middleware approach for the integration of airborne/ground Air Traffic Management systems for SOA-based future aeronautical service integration. Analysis of airborne/ground information exchange targeting on the domain-based operations and interactions are defined according to the added-value services, data access services and technical services in the SWIM context. An architectural overview of the SANDRA Airborne Middleware is illustrated featuring the connections between the onboard middleware and the SWIM Ground Gateways with its underlying system infrastructure. The "Delta" concept and compression mechanisms are proposed to optimise system performance and message exchange based on the selected technologies.

6. Acknowledgment

n° 233679. The SANDRA project is a Large Scale Integrating Project for the FP7 Topic AAT.2008.4.4.2 (Integrated approach to network centric aircraft communications for global aircraft operations). The project has 31 partners and started on 1st October 2009.

In addition, as an acknowledgement from the authors, the chapter represents only the opinion of the authors but not the whole SANDRA Consortium.

7. References

AIXM. (2011) AIXM Aeronautical Information Exchange Model, 12.12.2010, Available from http://www.aixm.aero/public/subsite_homepage/homepage.html

Andrews, T.; Curbera, F.; Dholakia, H.; Goland, Y.; Klein, J.; Leymann, F.; Liu, K.; Roller, D.; Smith, D.; Thatte, S.; Trickovic,I. & Weerawarana, S. (5th May 2003). Business Process Execution Language for Web Services Version 1.1, 24.11.2010, Available from http://xml.coverpages.org/BPELv11-May052003Final.pdf

Campbell, A.T.; Coulson, G. & Kounavis, M.E. (1999). Managing Complexity: Middleware Explained, *IEEE Computer Society IT Professional*, vol. 1, September | October 1999, pp. 22-28.

EasyJet & Ryanair. (2009). *Ryanair and EasyJet continue strong growth, British Airways' struggles continue in Oct-2009*.

Erl, T. (2009). *SOA design patterns (Ed. 1)*, Prentice Hall PTR, ISBN: 9780136135166, New Jersey, USA.

EUROCAE. (2006). *D7 Flight Object Interoperability Proposed Standard (FOIPS) Study. Issue 1.4.*

EUROCAE. (2009) *ED-133 Flight Object Interoperability Specification*.

EUROCONTROL. (2006). *Long-Term Forecast Flight Movements (2006 - 2025) v1.0*, EUROCONTROL.

EUROCONTROL. (2008). What is SESAR WP 14 SWIM technical architecture, 21.10.2010, Available from http://www.swim-suit.aero/swimsuit/project.php

EUROCONTROL. (2011). Weather Information Exchange Model, 12.12.2010, Available from http://www.eurocontrol.int/aim/public/standard_page/met_wie.html

FAA. (28th April 2010). SWIM Documents, 11.11.2010, Available from http://www.faa.gov/about/office_org/headquarters_offices/ato/service_units/t echops/atc_comms_services/swim/documentation/

GEIA. (March 2008) FAA *SWIM Program SOA Best Practices - Industry Input*

Houdebert, R. & Ayral, B. (2010). Making SWIM Interoperable between US and Europe, *Integrated Communications Navigation and Surveillance (ICNS) Conference Proceedings of IEEE*, Herndon, Virginia, USA, May 11-13, 2010.

ICAO. (2010). Operational Opportunity. *ICAO Environmental Report 2010*, Chapter 3, pp.96-125

Luckenbaugh, G.; Landriau. S.; Dehn, J. & Rudolph, S. (2007). Service Oriented Architecture for the Next Generation Air Transportation System, *Integrated Communications Navigation and Surveillance (ICNS) Conference Proceedings of IEEE*, Herndon, Virginia, USA, May 1-3, 2007.

Mahmoud, Q. H. (2004) *Middleware for Communications (Ed. 1)*, John Wiley & Sons, Inc., ISBN: 0470862068, New York, USA.

SANDRA. (June 2010) *Deliverable D3.1.1 SANDRA Detailed Network Requirements v1.0.*

Simon, R. (2003) Middleware, In: *The Internet Encyclopedia*, Bidgoli, H., vol. 1, 603-613, John Wiley & Sons, Inc. ISBN: 978-0-471-22202-6, New Jersey, USA.

SWIM-SUIT. (June 2008) *Deliverable D1.5.1 Overall SWIM Users Requirements v01.*

W3C. (2004) Web Services Architecture, 11.02.2004, Available from http://www.w3.org/TR/ws-arch/

Transport Protocol for Future Aeronautics

Muhammad Muhammad and Matteo Berioli
German Aerospace Center (DLR), Oberpfaffenhofen
Germany

1. Introduction

Transmission Control Protocol (TCP) (Postel, 1981) is the vastly researched, used and implemented transport protocol from the bouquet of standardized protocols we have. Since its first introduction, TCP has experienced tremendous revolutions specifically regarding the way it should control congestion (Allman et al., 2009). Many TCP flavors exist, one can control its sending rate according to link bandwidth estimation like in the Westwood variant (Mascolo et al., 2001), while another tunes its congestion window (cwnd) based on the round-trip time (RTT) evaluation such as TCP Vegas (Brakmo & Peterson, 1995), just to name few examples. However, all of these TCP versions assume a bulk transfer of data and thus they are optimized upon this type of traffic.

On the other hand, Air Traffic Management (ATM) data traffic has other characteristics than file transfer. Messages generated from aeronautical services are triggered by events in the aeronautical environment. Further, these messages have bursty inter-arrival times, which can go up to several minutes, relatively small size mostly in the order of few hundreds of bytes and a maximum delay (a maximum of few seconds) that they have to be delivered within. As can be realized, a TCP source may experience some inactive periods due to the burstiness of the traffic. However, in (Jacobson, 1988), it is recommended that to control and avoid congestion, TCP should use slow start mechanism when restarting a transmission after a ratherly long idle period.

The NEWSKY project (NEWSKY, 2009) recommends to design a transport protocol for avionics which is based on the User Datagram Protocol (UDP) but with reliability measures. Taking these suggestions into account, the Reliable User Datagram Protocol (RUDP) (Bova & Krivoruchka, 1999) is an interesting option. However, it has features from TCP, which are not desirable for the ATM context, like connection initiation and shutdown, congestion control and byte streaming mechanism, which according to the authors of (NEWSKY, 2009) and (Muhammad & Berioli, 2010) reduce the performance of a transport protocol when transferring small files. Therefore, RUDP could be a highly qualified candidate to transport aeronautical traffic in case it is carefully re-designed and adjusted to consider the properties of an ATM traffic over highly asymmetrical links.

Furthermore, in the context of checking the validity of TCP to transfer messages of ATM services, questions related to the performance of the protocol will be addressed. For example, the relation between the *cwnd* of TCP, aeronautical messages and congestion control.

The Seamless Aeronautical Networking through integration of Data links, Radios and Antennas (SANDRA) (SANDRA, 2011) concept consists of the integration of complex and disparate communication media into a lean and coherent architecture. SANDRA as a project

is divided into several sub projects (SPs). This work belongs to the one dealing with seamless networking specifically for the transport layer protocols and performance task. Transport layer protocols should be assessed with respect to their suitability for aeronautical communication and their impact on the system availability and reliability. Finally, an optimization for these protocols in order to be used in avionics context should be provided.

The work in this chapter is organized as follows. In the next Section, an overview about the aeronautical services is presented. In Section 3, the system architecture design in which the protocol will be operating is discussed. Section 4 explains some properties and illustrates few drawbacks of using TCP in avionics from theoretical and technical perspectives. The RUDP is further investigated in Section 5 where adjustment recommendations are given. In Section 6, detailed analysis of the ATM traffic pattern is shown and suggestions on system design are provided. Finally, a conclusion is drawn in Section 7.

2. Aeronautical services

The currently operating Aeronautical Telecommunication Network (ATN) protocol is based on the ISO/OSI stack and uses the TP4 protocol (ITU-T, 1995) via Dialogue Service (DS). However, the future envisaged protocol is expected to be based on the IPv4/v6 protocol stack, as identified in (NEWSKY, 2009).

The study of potential future communications technologies to meet safety and regularity of flight communications requirements, i.e. those supporting Air Traffic Services (ATS) and Airline Operational Control (AOC), have been initiated by the European Organization for the Safety of Air Navigation (EUROCONTROL) and the Federal Aviation Administration (FAA). The second version of the Communications Operating Concept and Requirements (COCR) document (EUROCONTROL & FAA, 2007) identifies the requirements placed on the communications that take place between the aircraft and ground radios, i.e. the air-to-ground (A/G) and ground-to-air (G/A) links. Further, the COCR divides the airspace into five domains. Thus, ATM services vary according to the domain in which the position of the airplane will be. Figure 1 illustrates the airspace domains, which are:

- Airport (APT) consists of the airport and the close area surrounding the airport.
- Terminal Maneuvering Area (TMA) also surrounding the airport but on a larger scale.
- En Route (ENR) is the continental or domestic airspace used by Air Traffic Control (ATC).
- Oceanic, Remote, Polar (ORP) is the same as ENR but it covers geographical areas generally outside the domestic airspace.
- Autonomous Operations Area (AOA), in this block of airspace, the airplane is self-separate, i.e. ATC is not used.

The corner stone of this work is to introduce a reliable communication environment for the ATS and AOC services. The ATS applications are the most critical for the safety of the flight. They involve the messages between the cockpit and the control center, i.e. the ATC center in the airport. Paired to these services are the AOC, which are primarily concerned with the safety and regularity of the flight. AOC applications are responsible for the communication between the aircraft and the AOC center, company or operational staff at an airport.

The messages generated by these applications have inter-arrival times in the order of seconds up to several minutes. Moreover, each message has a maximum latency time that it has to be delivered within. Typically, this time is lower than a message inter-arrival time. It is also worth to mention that the ATM traffic is inelastic; this means that these services cannot adjust their
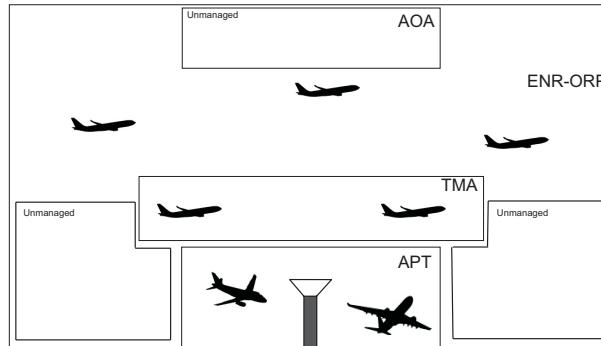
Fig. 1. Airspace domains as envisaged by the COCR document, namely, APT, TMA, ENR, ORP and AOA

message generation to the network conditions but they require a minimum level of resources, see (EUROCONTROL & FAA, 2007).

Within the COCR document, there are 32 ATS services, 21 belong to AOC and 2 NET services. The messages generated by these services have relatively small size with respect to a common file on the Internet. For example, the *FREETEXT* message is only 377 *bytes*. This service represents a textual message between the cockpit and the AOC center. Very few messages are of kilobytes size. The shortest message size is 82 *bytes* while the largest is around 21 *kilobytes*, which deals with the weather reports and is called *WXGRAPH*. Section 6 will present detailed analysis about the aeronautical traffic pattern.

3. System architecture

The trend in the aeronautical sector, driven by an increasing number of flights, is the introduction of new communication services and the transition of the ATM from analogue techniques towards digital communication services results in the development and deployment of new and heterogeneous link technologies, see (Kissling, 2009). One example of a link technology under development for direct A/G radio communication is the L-band Digital Aeronautical Communication System (L-DACS)-1 communication standard (Brandes et al., 2009) and (Graupl et al., 2009). Another example is the European Space Agency (ESA) Iris programme (ESA, 2009) using satellite communications, which is currently under development. The new link technologies will complement the existing ones (like VHF Digital Model (VDL)-x) and will also coexist with future, short range links such as WiMAX as part of the communication infrastructure in the APT area. The system architecture under consideration in this work requires the aircraft to have several link technologies available for communication with the ground, which all have heterogeneous properties. One of the most important properties changing with the link is the propagation delay with RTTs in the order of 500 *ms* for GEO satellites down to few milliseconds for direct A/G links. Besides the delay, the communication bandwidth and datarates change as well with the links, starting with few bits per second and reaching higher datarates with kilobits per second. Finally, also from the channel characteristics, packet error and loss rates may change with the link technology. The integration of these different links into one network results in a system architecture as illustrated (simplified) in Figure 2. As shown there, only the ATS and AOC services are considered. Each of these services may have one or
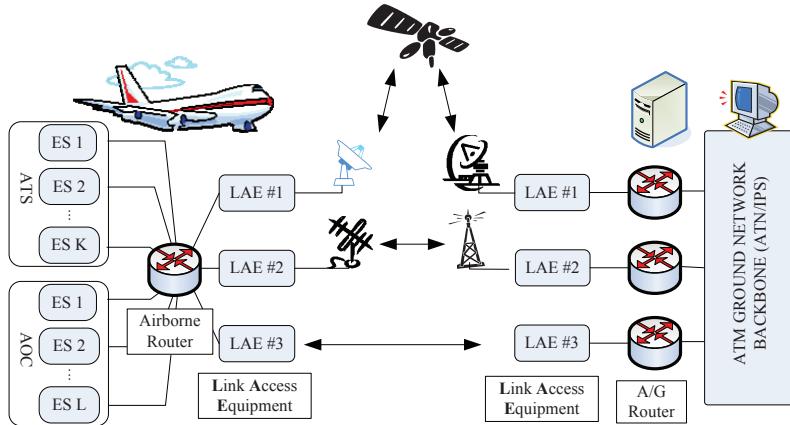
Fig. 2. System architecture

several End Systems (ESs), which generate and receive data messages. A data message is transmitted via the airborne router as one or several IP packets over a link with a specific link layer and physical layer protocols belonging to the selected Link Access Equipment (LAE). On the ground side, the network layer IP datagrams are then routed via the ATM ground network (assumed as a backbone ATN/IPS network) towards the correspondent ES in the ground network. For safety related operational services ATS and AOC, reliable message transmission is a key demand since loss or corruption of messages can have severe consequences. Transport layer protocols, which shall guarantee the reliable transmission of data should work efficiently over all different link technologies present. For provision of transmission reliability in the future ATN/IPS network, up to now the TCP protocol has been envisaged (ICAO, 2010). The different conceptual options to use the TCP protocol and the requirements to a transport protocol have been reported in (Kissling & Graupl, 2008) and (Ehammer, Graupl, Rokitansky & Kissling, 2009). In (Kissling & Baudoin, 2008), different possibilities for protocol stack architectures in the ATM scenario have been investigated. The major key issues for using transport layer protocols in the ATM environment found in this work were the provision of a reliable communication service which makes best possible use of the scarce communication bandwidth and does not require coexistence of different transport layer protocols in parallel. Deployment of several link-specialized transport protocols in parallel would cause higher cost for maintenance, management and especially for the standardization procedure which every transport layer protocol has to go through.

4. Transmission control protocol

TCP is the commonly used protocol by many of today's Internet most popular applications like Hypertext Transfer Protocol (HTTP) and File Transfer Protocol (FTP). It provides reliability and ordered delivery of a stream of bytes being transmitted from a program on one computer to another over an unreliable network.

TCP is a connection oriented protocol. That is, a connection should be established before sending data. Depending on the RTT between the two communicating nodes, a waiting time of:

$$T_{\mathrm{w}} = 2 \cdot RTT \tag{1}$$

is measured before receiving the first byte of any actual information.

The work done by (Ehammer, Graupl, Rokitansky & Kissling, 2009) describes several possibilities on the way that TCP could be operated over aeronautical networks and showed the advantages and drawbacks of each method. These methods deal with the multiplexing or not of the ATM services and the network layer interaction. These techniques are: (a) re-establishing of connections for each transmission and for every service (Figure 3(a)), (b) establishing a connection once for each service and keep it open (Figure 3(b)), (c) re-establishing of the connection for each transmission of a multiplexed set of services (Figure 4(a)) and (d) establishing a connection for a multiplexed set of services and keep it open (Figure 4(b)).

(a) Connection established/shutdown for every transmission

(b) Connection established/shutdown once for every service
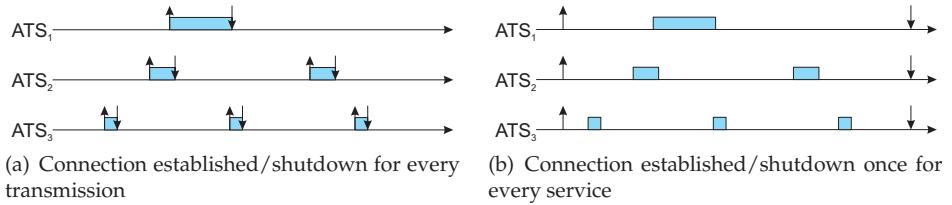
Fig. 3. Transport layer session connection establishment/shutdown with no services multiplexing

(a) Connection established/shutdown for every multiplexed transmission

(b) Connection established/shutdown once for all multiplexed services
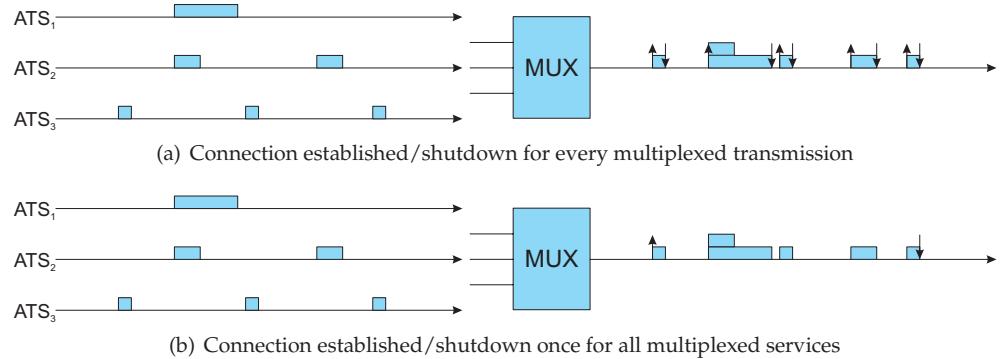
Fig. 4. Transport layer session connection establishment/shutdown with services multiplexing

Further, in (Ehammer, Graupl & Rokitansky, 2009), the authors discussed several well known TCP optimization techniques and found out that for TCP to operate over aeronautical communication network, a minimum of a certain link capacity should be available per user, else TCP will run into performance problems.

Moreover, within the aeronautical environment, where the messages to be transmitted in a flight are scarce, maintaining a single TCP connection is cumbersome. This is due to the fact that within a flight the aircraft will change various links with different capacities and delays, giving rise to links handover management overhead. In (Daniel & Kojo, 2008) and (Daniel et al., 2008) the authors had developed cross-layer assisted TCP sender algorithms

to reduce the unnecessary packet retransmissions and congestion control actions due to vertical handovers. In other words, they developed mechanisms to help TCP to reduce retransmission overhead when operated in wireless environment, in which handovers may occur on events such as the links properties are interchanged between high/low delays, high/low bandwidth or upon retransmission timeouts (RTOs) during disconnection.

On the other hand, allowing TCP to establish a session for every transmission brings three disadvantages. Figure 5 illustrates this approach in comparison with the single TCP connection (discussed previously). First, for every message to be received, T_w should be waited before the first bit of the actual data delivery, which affects the delivery time of the message. Secondly, the overhead introduced by TCP, because of the connection initiation and shutdown is on average large compared to the sizes of the messages produced by the ATM services. Finally, the *cwnd* of a TCP sender will not be able to capture the congestion status when operating below aeronautical services because of the traffic pattern they adhere.
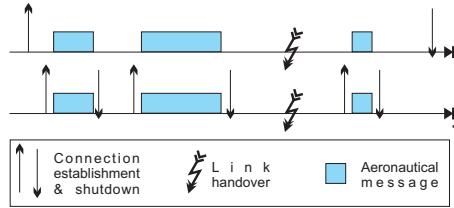
Fig. 5. Transport layer (TCP) connections as envisaged in (c) and (d)

In (Muhammad & Berioli, 2010), we developed a model to theoretically estimate the overhead produced by different transport protocols when transferring a file. The investigations showed that for small file sizes, reliable, connectionless transport protocols with no congestion and flow control have the lowest overhead cost. In this work, we extend the model to measure the overhead produced by the connection establishment and shutdown mechanisms of TCP. According to the specification of TCP in (Postel, 1981), to start a session, a client should send a SYN packet to the server that will respond with a $SYN - ACK$ and finally the client will confirm with an ACK (3-way handshake). On the other hand, when the server finishes transferring the file, it will send a FIN packet, and the client will reply with an ACK. Further, when the client makes sure that it has the complete contents of the file, it will send another FIN to the sender and wait for the ACK packet. These two mechanisms are bi-directional and thus we split our model based on the forward (FW) and return (RT) channels, respectively.

$$\begin{aligned} S_{HC}^{FW} &= H_{SYN-ACK} + H_{FIN} + H_{ACK}, \\ S_{HC}^{RT} &= H_{SYN} + H_{FIN} + 2 \cdot H_{ACK} \end{aligned} \tag{2}$$

where S_{HC}^{FW} and S_{HC}^{RT} are the sizes of the overhead on the FW and RT channels, respectively. Moreover, H_X is the size of the header of the packet X, bearing in mind that TCP connection establishment and shutdown packets are only TCP headers.

Now, the overall overhead related to a connection start and end in TCP amounts for:

$$H_C = S_{HC}^{FW} + S_{HC}^{RT} \tag{3}$$

Therefore, the TCP session overhead produced for a file of size F is:

$$OH_C = \frac{H_C}{F} \tag{4}$$

(a) One message per transmission   (b) Multiple connections
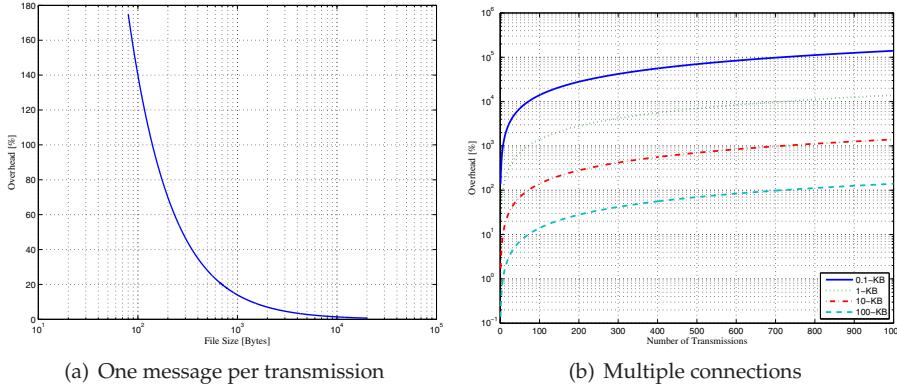
Fig. 6. Overhead estimation of a TCP connection establishment and shutdown

The two plots in Figure 6 show the overhead generated for different message sizes and the overhead cost for transmitting the same file multiple times with different TCP connections. For this simulation, we used the minimum TCP header size (20 *bytes*) assuming no options are used. In Figure 6(a), we show the connection overhead percentage (on the *y-axis*) generated for different sizes of messages in the range of the ATM services messages size (*x-axis*). Here, we assume a single message is transferred in a connection. As can be clearly seen that the overhead of sending a message of small size is considerably large.

Moreover, Figure 6(b) represents the overhead generated for multiple file sizes with several TCP connections. For instance, when simulating a file of size 10 *kilobytes*, the *x-axis* represents the number of transmissions of that file each time with a new connection initiated and shutdown. As it can be realized, the higher the number of transmissions, the higher the overhead. Conversely, the larger the file, the lower the overhead.

Taking the complete traffic profile characteristics into account and not a single message as above, Figure 7 shows the *cwnd* (displayed with a *line*), in bytes, of TCP versus the actual transmitted bytes (represented by *stars*). It can be clearly realized that the *cwnd* of TCP cannot cope with the real transmitted information. We can say that the *cwnd* of TCP was deceived by the incoming application data. Thus, it shows a very oscillating behavior and this is related to the long idle periods between two consecutive transmissions, which is associated with the inter-arrival times of the ATM messages. Further, the average size of this *cwnd* (reflected by the *dotted line*) is very close to the value of the average *cwnd* resulted in the simulations done by (Ehammer, Graupl, Rokitansky & Kissling, 2009).

5. Candidate protocols

In the summer of 1984, the standardization of the Reliable Data Protocol (RDP) was held by (Velten et al., 1984) with the idea of running applications such as remote loading and debugging. Six years later, their colleagues at BBN Communications Corp. updated the standard with a newer version (Partridge & Hinden, 1990). In short, RDP is connection oriented and offers reliable transport to efficiently support the bulk transfer of data.

Moreover, based on these two standards with the same protocol properties, RUDP defined in (Bova & Krivoruchka, 1999) (IETF Internet-draft) extended the primitive[1] UDP (Postel,

---

[1] Because, according to the authors, TCP is too complex.
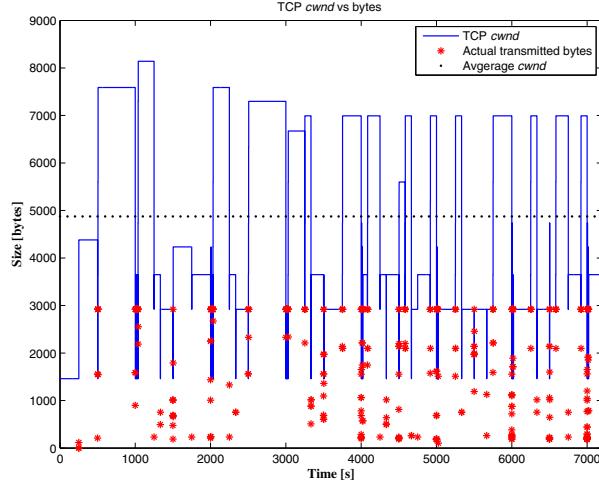
Fig. 7. *cwnd* of TCP vs the actual transmitted bytes from the aeronautical services

1980) with window and flow control, acknowledgement mechanism, retransmission of packets lost on the way and avoidance of buffer overflow in order to transport telephony signaling.

As mentioned in Section 2, each aeronautical service sends a different message. The generation of these messages is triggered by events. The services at the two ESs, i.e. the airplane and the control center, send their messages independently. Further, the generated messages differ significantly among each other in terms of message arrival time, size and latency requirements. However, almost all of them require reliable delivery service by the transport protocol. Some services do not require full reliability like surveillance ones because they are generated more frequently, see (Ehammer, Graupl & Rokitansky, 2009). Thus, a reliable transport protocol to be designed should be aware of the properties of these messages. Figure 8 illustrates the protocol stack using the provisioned transport protocol above the Internet Protocol (IP) layer that takes care of the addressing and below the aeronautical services labeled $A_1, A_2, ...$ An aeronautical transport protocol should provide reliability, ordered delivery and honors message boundaries like UDP. Further, it should be IP based and connectionless transport protocol in order to reduce the session initiation and shutdown overhead. Finally, it should be able to cope with highly asymmetrical communication channels such as satellite links.
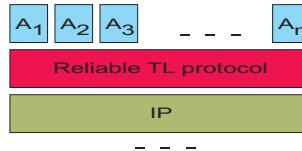
Fig. 8. The protocol stack to be used by a reliable transport UDP-based protocol such as enhanced version of RUDP

Aeronautical communications involve layer 2 (Media Access Control (MAC) layer) congestion control mechanisms such as L-DACS. Therefore, the motivation behind removing the

congestion control methods from the transport layer (layer 4) is to prevent any further decrease of the performance in case the two layers mechanisms clash. Further, in this case, it makes more sense to move the congestion control to the MAC layer since they have the knowledge about the links properties and thus there is no need for cross-layer communication. Therefore, reducing system complexity.

Furthermore, in (Ehammer, Graupl & Rokitansky, 2009), it was shown that the *WXGRAPH* message on the FW link (sized 21 *kilobytes*) is the one that is congesting the network due to its large size compared to the other services. One proposal could be, the implementation of a dual stack. Where TCP takes care of transmitting the messages from this service and keeps the others of small sizes for the new protocol.

6. System dimensioning

The requirements of the aeronautical traffic mandate a certain behavior from the lower layers. This behavior is governed by the properties offered by the system. This section will elaborate more on the aeronautical traffic properties. Additionally, some recommendations on the system properties will be derived.

6.1 Aeronautical traffic properties

As described in Section 2, messages generated by aeronautical services vary in size, inter-arrival times and latency requirements. These variations do not only occur between two different airspace domains or links (A/G and G/A links) but also within the same transport layer connection in case all services are multiplexed on the same transport session, as illustrated in Figure 4(b), which is an option from the transport layer session management mechanisms highlighted in Section 4. Further, Figures 9 and 10 show an aeronautical traffic pattern on the FW and the RT links, respectively. These plots show a sample test for two hours flight simulation in the TMA-ENR domain based on statistical expected traffic reports from (Rokitansky et al., 2008).

Figures 9(a) and 10(a) show the size of the messages and their inter-arrival times from the application layer perspective. It is clear to see that the size of the messages fluctuates heavily especially on the FW link; on the RT link the largest measured message size is much smaller than the one on the FW channel; the inter-arrival times of these messages vary not only among the different services but also within every service as well. Nevertheless, few services are generated periodically.

Figures 9(b) and 10(b) represent the data flows of these aeronautical messages. These data flows can be described by means of the cumulative function $D(t)$, defined as the number of bits seen on the flow in the time interval $[0, t]$:

$$D(t) = \sum_i l_i \tag{5}$$

where l_i is the length in bytes of the message i.

The plots shown in Figure 11 represent the amount of bytes related to a specific service - identified by its size on the *x-axis*. For instance, the upper plot tells us that 80 % of the traffic data on the FW channel is generated by the *WXGRAPH* service (in which a single message amounts to 21 *kilobytes*). Additionally, all other services generate the remaining 20 % of the data traffic. In other words, at the network layer this can be seen as a flow of packets with more than 80 % of large sized datagrams and ca. 12 % of packets less than the maximum transmission unit (MTU) size of 1500 *bytes*.

(a) Messages generated on the FW link

(b) The cumulative function $D(t)$ on the FW link
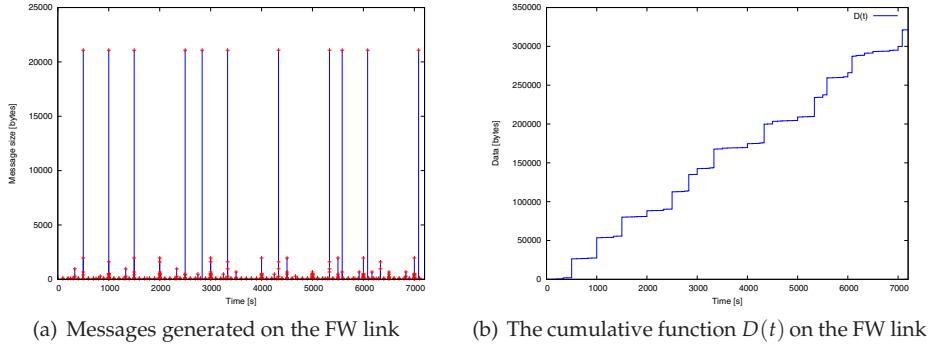
Fig. 9. Aeronautical traffic flow on the FW link in the TMA-ENR domain

(a) Messages generated on the RT link

(b) The cumulative function $D(t)$ on the RT link
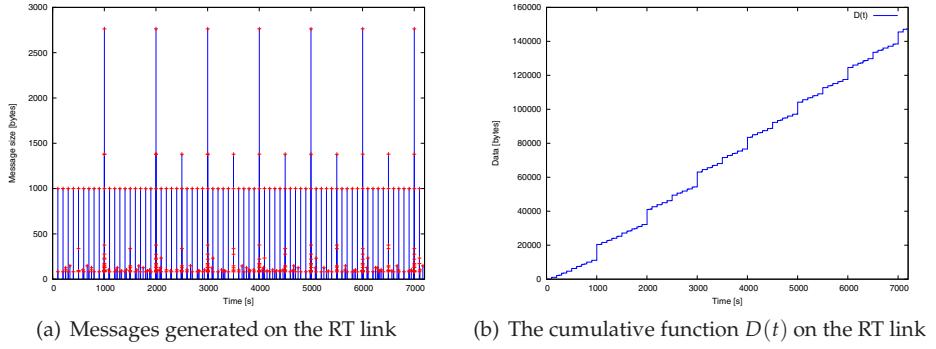
Fig. 10. Aeronautical traffic flow on the RT link in the TMA-ENR domain

On the other hand, only a single aeronautical service on the RT channel generates a message of size greater than the MTU size. However, this message can be fragmented only in two IP datagrams, if we also consider an MTU size similar to the one on the FW channel. Further, one can realize from the chart that almost 60 % of the traffic on the RT channel is generated by a service message of size 1000 *bytes*.

6.2 System properties

One of the major differences between an aeronautical application and a file transfer application, like FTP, is that avionics services are *inelastic*. This means that a service requires a certain minimum level of bandwidth and a certain maximum latency. In other words, an aeronautical application cannot adjust its rate, for example, to changeable network conditions. By contrast, an elastic application can adapt to network conditions. A file transfer, for instance, can easily adjust to different level of available bandwidth and latency as it has no severe latency requirements.

Based on the above mentioned facts and in order for the system to function properly, it is fundamental to define some minimum system requirements such as serving rate and queues length. Figure 12 illustrates a simplified network architecture in which the ATS/AOC client messages are buffered in a queue on the mobile router. This set-up is implemented on every
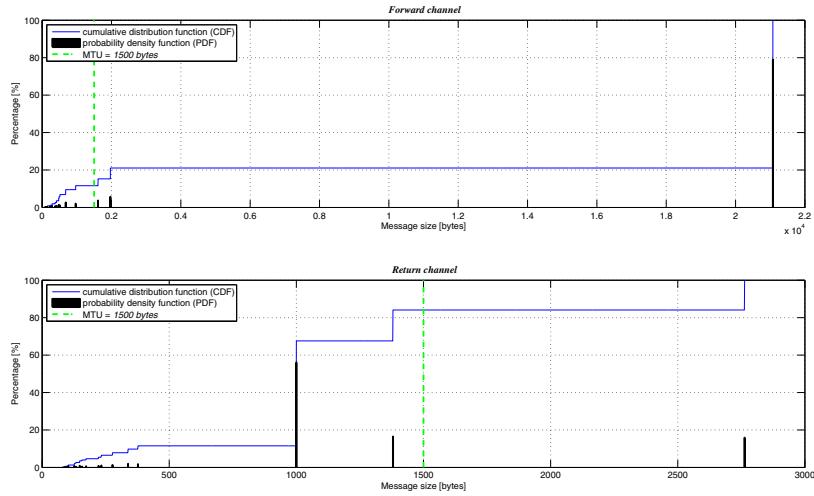
Fig. 11. Amount of data traffic generated by different ATM messages on the FW and RT channels, respectively

aircraft. Access to the wireless link to connect to the ground station is granted through layer 2, in the protocol stack (MAC layer). At the other end, i.e. the ground segment, the corresponding servers use a single queue, at the ground station, to buffer the messages directed to all airplanes.
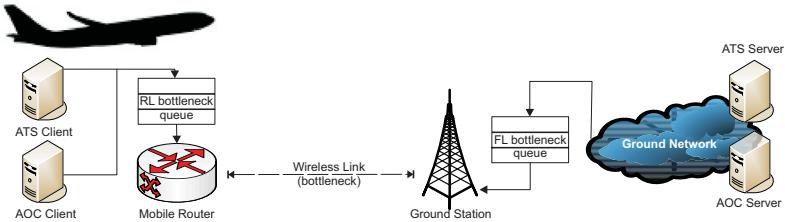
Fig. 12. Simplified network architecture

To illustrate this, let's assume that we have on the RT channel the service *WXGRAPH*, which is transmitting messages of a fixed size equal to 93 *bytes* with inter-arrival times illustrated in Figure 13(a). Figure 13(b), on the other hand, represents a possible realization of $D(t)$, as defined in (5).

Messages or packets (of fragmented messages) arrive (with a cumulative amount of data $D(t)$) at a buffer of size B and they are served in a First In First Out (FIFO) order with a rate r. The serving rate r exhibits the time needed for a message of size l to be transmitted on the link, in general:

$$t = \frac{l}{r} \tag{6}$$

This scenario can be easily modeled using a leaky bucket as shown in Figure 14. A leaky bucket controller, according to (Boudec & Thiran, 2001), is a device that handles the data flow
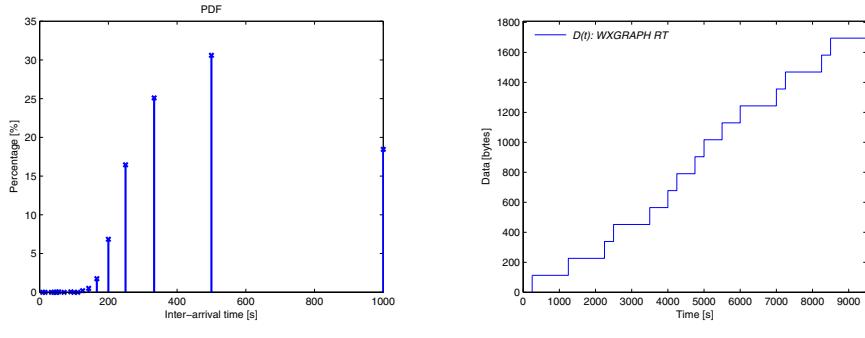
(a) PDF of the inter-arrival times for the *WXGRAPH* service on the RT link

(b) Cumulative function $D(t)$ of the *WXGRAPH* service on the RT link

Fig. 13. Properties of the *WXGRAPH* service on the RT link as reported in (Rokitansky et al., 2008)

$D(t)$ as follows. There is bucket of fluid of size B. The bucket is initially empty ($t = 0$). The bucket has a hole and leaks at rate of r units of fluid per second when it is not empty. The data flow $D(t)$ pours into the bucket in the time interval $[t_0, t]$ an amount of fluid equal to the amount of data $D(t) - D(t_0)$. Data that would cause the bucket to overflow is discarded.
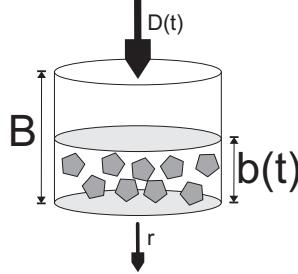
Fig. 14. Illustration of a leaky bucket model, pentagons represent the conformant data

6.2.1 Serving rate: r

Before deriving the minimum value of the required serving rate r, we need to make some assumptions such that the following work is clearer. First, let G_ζ describe a traffic generation stochastic process with emission rate λ_ζ, which is a stochastic variable with $\bar{\lambda}_\zeta = \mathrm{E}\left[\lambda_\zeta\right] = \frac{1}{\bar{t}_\zeta}$, where $\bar{t}_\zeta$ is the average inter-arrival time. Further, assume that the emitted traffic is reaching the queue (the object to be studied) with the same rate, λ_ζ. So, λ_ζ is also an arrival rate process. Then, we denote by:

$$G = \sum_\zeta G_\zeta \tag{7}$$

the total processes with an aggregate generation rate of:

$$\lambda = \sum_\zeta \lambda_\zeta \tag{8}$$

where $\zeta \in \Phi$, which is the number of processes. This follows:

$$\bar{\lambda} = \sum_\zeta \bar{\lambda}_\zeta \tag{9}$$

where $\bar{\lambda}$ is the average aggregate arrival rate of the process G, and the average inter-arrival time is:

$$\begin{aligned} \bar{t} &= \frac{1}{\bar{\lambda}} \\ &= \frac{1}{\sum_\zeta \bar{\lambda}_\zeta} \\ &= \frac{1}{\sum_\zeta \frac{1}{\bar{t}_\zeta}} \end{aligned} \tag{10}$$

At the first glance, we are interested in the behavior of a single process/service G_ζ, as if it is being served by a FIFO queue with serving rate r_ζ.

For a particular service G_ζ, if the number of arrivals in an interval $[0, t]$ is $\alpha_\zeta(t)$, then we can define:

$$\lambda_\zeta(t) \triangleq \frac{\alpha_\zeta(t)}{t} \tag{11}$$

as the arrival rate of the service ζ. Consequently, the average arrival rate and inter-arrival time of the service can be deduced from the following; by assuming that this limit exists:

$$\begin{aligned} \bar{\lambda}_\zeta &= \lim_{t \to +\infty} \lambda_\zeta(t) \\ &= \frac{1}{\bar{t}_\zeta} \end{aligned} \tag{12}$$

The service rate r_ζ allows us to determine the time needed for a service message of size l_ζ to be transmitted on the link, which is specified in (6). As illustrated in Figure 15(a), if the serving rate is low, then the upcoming messages will be buffered, until the current ones are processed, and therefore all message are delayed.

Now, given the fixed message length l_ζ of a service ζ and the probability density function (PDF) of the inter-arrival times, we can determine the minimum serving rate required for this service. The minimum required serving rate r_ζ for service ζ can be evaluated by:

$$\begin{aligned} r_\zeta &= \frac{l_\zeta}{\bar{t}_\zeta} \\ &= l_\zeta \cdot \bar{\lambda}_\zeta \end{aligned} \tag{13}$$

From Figure 15(b), this can be understood considering that r_ζ is the *average slope* of the function $D_\zeta(t)$; it is clear that if the flow $D_\zeta(t)$ is served at a rate lower than r_ζ, the distance between $D_\zeta(t)$ and the dotted line (offered load) will increase indefinitely for $t \to +\infty$ (see "waiting time" in Figure 15(a)). Thus, this distance remains bounded only if the serving rate is higher or equal to r_ζ. In this case every message is being served the moment it enters the bucket; sometimes a message has to wait for some time in case the server is processing a previous job. This small waiting time is due to the fact that r_ζ is derived based on the average inter-arrival time, therefore, it represents a lower bound on the serving rate for a single service ζ.
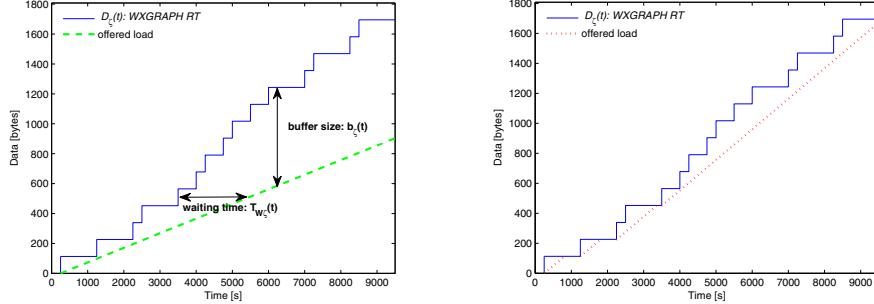
(a) $D_\zeta(t)$ with low serving rate r_ζ; G_ζ is WXGRAPH-RT

(b) $D_\zeta(t)$ with the minimum required serving rate r_ζ; G_ζ is WXGRAPH-RT

Fig. 15. Different serving rates for the $D_\zeta(t)$ of WXGRAPH-RT

The estimation of the minimum required rate for an aggregation G of traffic sources can be similarly analyzed. If the total number of arrivals of the G service in the interval $[0, t]$ is denoted by $\alpha(t) \triangleq \sum_\zeta \alpha_\zeta(t)$, then from (11), the total arrival rate is:

$$
\begin{aligned}
\lambda(t) &\triangleq \frac{\alpha(t)}{t} \\
&= \frac{\sum_\zeta \alpha_\zeta(t)}{t} \\
&= \sum_\zeta \lambda_\zeta(t)
\end{aligned}
\tag{14}
$$

in which, if we assume that for every $\zeta \in \Phi$ all limits as in (12) exist, as t approaches $+\infty$, we obtain the average of the aggregate arrival rate:

$$
\begin{aligned}
\bar{\lambda} &= \lim_{t \to +\infty} \lambda(t) \\
&= \lim_{t \to +\infty} \sum_\zeta \lambda_\zeta(t) \\
&= \sum_\zeta \lim_{t \to +\infty} \lambda_\zeta(t) \\
&= \sum_\zeta \bar{\lambda}_\zeta(t)
\end{aligned}
\tag{15}
$$

Now, the ratio of the number of arrivals of service ζ to the total arrivals can be deduced from:

$$
\frac{\alpha_\zeta(t)}{\alpha(t)} = \frac{\lambda_\zeta(t)}{\lambda(t)}
$$

which as $t \to +\infty$ converges to

$$
= \frac{\bar{\lambda}_\zeta}{\bar{\lambda}}
\tag{16}
$$

It follows that we can denote the average message size $\bar{l}$ for the aggregate process G as:

$$\bar{l} = \sum_{\zeta} \frac{\bar{\lambda}_{\zeta}}{\bar{\lambda}} \cdot l_{\zeta} \tag{17}$$

which is basically dependent on the average arrival rate of every service.

Finally, the overall minimum required serving rate r for the aggregation of all the services $\zeta \in \Phi$ is:

$$
\begin{aligned}
r &= \bar{\lambda} \cdot \bar{l} \\
&= \bar{\lambda} \cdot \sum_{\zeta} \frac{\bar{\lambda}_{\zeta}}{\bar{\lambda}} \cdot l_{\zeta} \\
&= \sum_{\zeta} \bar{\lambda}_{\zeta} \cdot l_{\zeta} \\
&= \sum_{\zeta} r_{\zeta}
\end{aligned} \tag{18}
$$

which is simply the aggregate serving rate of all services $\zeta \in \Phi$.

6.2.2 Buffer size: B

The buffer size B plays a vital role in aeronautical communications. It determines, besides the serving rate, the maximum waiting time that can be introduced by a queue also the probability to drop packets because of an overflow, i.e. the dropping probability. If B is very large, then a message/packet will experience a long waiting time that will affect the latency requirements of the service but a low dropping probability. Conversely, a small B will result in packets/messages being dropped whenever the buffer is full; thus, a higher dropping probability, but lower waiting time.

Also, starting the evaluation for a single process served at rate r_{ζ}, the waiting time of a packet, arriving at time t, in the queue is:

$$T_{W_{\zeta}}(t) = \frac{b_{\zeta}(t)}{r_{\zeta}} \tag{19}$$

where $b_{\zeta}(t)$ is the queue size at time t, i.e. $D_{\zeta}(t) - r_{\zeta} \cdot t$, see Figure 15(a).

Assuming that the limit as t approaches $+\infty$ exists for every service $\zeta \in \Phi$, we can determine the expected waiting time per service message in the buffer as:

$$\bar{T}_{W_{\zeta}} = \lim_{t \to +\infty} \frac{1}{t} \int_0^t T_{W_{\zeta}}(x)\,dx \tag{20}$$

which is independent of the used queueing model. Further, the average queue length for service ζ can be represented by:

$$\bar{b}_{\zeta} = \lim_{t \to +\infty} \frac{1}{t} \int_0^t b_{\zeta}(x)\,dx \tag{21}$$

If the last two limits in (20) and (21) exist, then the Little's theorem from (Kleinrock, 1975) states that the average buffer length in a queueing system can be rewritten as a function of the average process arrival rate times the average waiting time:

$$\bar{b}_{\zeta} = \bar{\lambda}_{\zeta} \cdot \bar{T}_{W_{\zeta}} \tag{22}$$

Equation (22) describes an average queue size for the service ζ. For this single service, the queueing system could be modeled as $M/D/1$ where M refers to memoryless arrival rate and D stands for deterministic serving rate since the for a single service the message size l_ζ is constant, which implies a constant serving time per message.

On the other hand, the aggregation of multiple services into a single queue allows the usage of a model with general serving time distribution, i.e. $M/G/1$, for which the size of the messages vary accordingly.

The above mentioned models describe a queue with a length that can grow indefinitely. Therefore, to limit it with a certain buffer size B, the queueing model $M/G/1/B$ could be implemented. This assumes that the aggregate arrival rate of the services has Poisson distribution, as proposed in the (EUROCONTROL & FAA, 2007). Clearly, for a finite buffer size B, the queue cannot accommodate all the messages arriving from the services because of their stochastic property. This is illustrated in Figure 16. As it can be clearly seen that not all of the total arrivals (λ) are being buffered since the queue can reject arrivals when the buffer is full. Thus, arrivals noted by (λ_e) denotes the rate of entering the system when the buffer is not full whereas (λ_d) indicates the rate of not entering the system because of the unavailability of resources.

Fig. 16. Illustration of a queue with limited size B and dropping probability

The work in literature about queueing systems, such as in (Kleinrock, 1975) and (Smith, 2004), allows the derivation of the dropping probability as a function of arrival rate, buffer size and serving rate. This helps in designing a system that takes the latency requirements of the services into consideration and with minimum losses.

7. Conclusions

The requirements of a simple but reliable transport layer protocol for the safety critical ATM services like ATS and AOC have been discussed. The basic assumption that draws the design phase of this study was based on a priori recommendations. The new protocol to be designed has to address two goals; first, take out some algorithms of TCP that complicate its operation and add overhead, specifically congestion and flow control plus session initiation and shutdown procedures. Secondly, cope with the pattern of the aeronautical traffic, which is different from the Internet traffic like FTP, for example. Within this study, some drawbacks of using TCP for aeronautical services were highlighted and some protocol enhancements for RUDP to be an alternative solution to overcome these weakness points of TCP have been recommended. Furthermore, the system properties in terms of service rate and buffer sizes on both ESs were discussed.

8. Acknowledgments

nř 233679.   The SANDRA project is a Large Scale Integrating Project for the FP7 Topic AAT.2008.4.4.2 (Integrated approach to network centric aircraft communications for global aircraft operations). The project has 31 partners and started on 1st October 2009.

9. References

Allman, M., Paxson, V. & Blanton, E. (2009). TCP Congestion Control. IETF RFC 5681.

Boudec, J.-Y. L. & Thiran, P. (2001). *Network Calculus: A Theory of Deterministic Queuing Systems for the Internet*, Springer-Verlag, Berlin, Heidelberg.

Bova, T. & Krivoruchka, T. (1999). Reliable User Datagram Protocol. IETF Internet-draft.

Brakmo, L. & Peterson, L. (1995). TCP Vegas: End to End Congestion Avoidance on a Global Internet, *IEEE Journal on selected Areas in communications* 13(8): 1465–1480.

Brandes, S., Epple, U., Gligorevic, S., Schnell, M., Haindl, B. & Sajatovic, M. (2009). Physical Layer Specification of the L-band Digital Aeronautical Communications System (L-DACS1), *Integrated Communications, Navigation and Surveillance Conference, 2009. ICNS '09.*, Arlington, Virginia, United States, pp. 1 –12.

Daniel, L., Järvinen, I. & Kojo, M. (2008). Combating Packet Reordering in Vertical Handoff Using Cross-Layer Notifications to TCP, *Proceedings of the 2008 IEEE International Conference on Wireless & Mobile Computing, Networking & Communication*, IEEE Computer Society, Washington, DC, USA, pp. 297–303.

Daniel, L. & Kojo, M. (2008). Employing Cross-Layer Assisted TCP Algorithms to Improve TCP Performance with Vertical Handoffs, *International Journal of Communication Networks and Distributed Systems* 1(4/5/6): 433–465.

Ehammer, M., Graupl, T. & Rokitansky, C.-H. (2009). TCP/IP over Aeronautical Communication Systems - Effects on Bandwidth Consumption, *IEEE/AIAA 28th Digital Avionics Systems Conference, 2009. DASC '09.*, FL, USA, pp. 4.A.3–1 –4.A.3–9.

Ehammer, M., Graupl, T., Rokitansky, C.-H. & Kissling, C. (2009). The Operation of TCP over Aeronautical Networks, *Integrated Communications, Navigation and Surveillance Conference, 2009. ICNS '09.*, Arlington, Virginia, United States, pp. 1 –11.

ESA (2009). ARTES 10 - Iris Programme.
     URL: *http://telecom.esa.int/iris*

EUROCONTROL & FAA (2007). Communications Operating Concept and Requirements for the Future Radio System, COCR v2.0.

Graupl, T., Ehammer, M. & Rokitansky, C.-H. (2009). L-DACS 1 Data Link Layer Design and Performance, *Integrated Communications, Navigation and Surveillance Conference, 2009. ICNS '09.*, Arlington, Virginia, United States, pp. 1 –12.

ICAO (2010). Manual for the ATN using IPS Standards and Protocols.

ITU-T (1995). Protocol for Providing the Connection-mode Transport Service, *Recommendation X.224*, International Telecommunication Union, Geneva.

Jacobson, V. (1988). Congestion Avoidance and Control, *SIGCOMM Comput. Commun. Rev.* 18(4): 314–329.

Kissling, C. (2009). Data Link Selection in Mobile Aeronautical Telecommunication Networks, *The 27th AIAA International Communication Satellite Systems Conference, ICSSC '09*, Heriot-Watt University, UK.

Kissling, C. & Baudoin, C. (2008). Protocol Stack Options in Heterogeneous Aeronautical Networks, *4th Advanced Satellite Mobile Systems, 2008. ASMS 2008*, Bologna, Italy, pp. 43 –48.

Kissling, C. & Graupl, T. (2008). Options for Using Transport Layer Protocols.

Kleinrock, L. (1975). *Queueing Systems Volume 1: Theory*, Wiley-Interscience.

Mascolo, S., Casetti, C., Gerla, M., Sanadidi, M. Y. & Wang, R. (2001). TCP Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links, *Proceedings of the 7th annual international conference on Mobile computing and networking*, MobiCom '01, ACM, New York, NY, USA, pp. 287–297.

Muhammad, M. & Berioli, M. (2010). Overhead Estimation of Selected Protocols for File Transfer, *The 5th Advanced satellite multimedia systems conference (asms)*, Cagliari, Italy, pp. 449 –454.

NEWSKY (2009). D12 Efficient Resource Management Techniques, Design, Assessment and Selection.

Partridge, C. & Hinden, R. (1990). Version 2 of the Reliable Data Protocol RDP. IETF RFC 1151.

Postel, J. (1980). User Datagram Protocol. IETF RFC 768.

Postel, J. (1981). Transmission Control Protocol. IETF RFC 793.

Rokitansky, C.-H., Ehammer, M. & Graupl, T. (2008). Communication Capacity Assesment for the Iris Satellite System, *IEEE/AIAA 27th Digital Avionics Systems Conference, 2008. DASC 2008*, FL, USA, pp. 2.B.3–1 –2.B.3–13.

SANDRA (2011). Seamless Aeronautical Networking through integration of Data links Radios and Antennas.
URL: *http://www.sandra.aero*

Velten, D., Hinden, R. & Sax, J. (1984). Reliable Data Protocol. IETF RFC 908.

Smith, J. M. (2004). Optimal Design and Performance Modelling of M/G/1/K Queueing Systems, *International Journal of Mathematical and Computer Modelling* 39(9/10): 1049–1081.

Security Concepts in IPv6 Based Aeronautical Communications

Tommaso Pecorella[1], Romano Fantacci[1], Luigia Micciullo[1],
Antonietta Stango[2], Neeli Prasad[2], Piotr Pacyna[3],
Norbert Rapacz[3] and Tomasz Chmielecki[3]

[1]*Università di Firenze, Firenze*
[2]*Center for TeleInFrastruktur (CTIF), Aalborg*
[3]*AGH University of Science and Technology, Kraków*
[1]*Italy*
[2]*Denmark*
[3]*Poland*

1. Introduction

Although aeronautical networks rely on communications, nowadays the largest part of such communications is based on old but proven standards, many of them being analogical and voice-based.

It is widely recognized that this communication model will not be able to support the increasing complexity and worldwide spread of aeronautical communications, especially with the ever-increasing number of players and locations.

The IP communication model seems a good candidate to replace the analogical communications. On the other hand, the debate about the well-known IPv4 pools depletion made it clear that the only viable and sustainable solution is to adopt IPv6 as a common basis. Concerning IPv4, it is believed that the only real need will be for passengers communications. IPv4 traffic can be segregated so to not harm the main network architecture.

IPv6 network security is not different, from a logical point of view, from IPv4 one. The main difference is related to a simple yet hard to fully admit concept: security is not subject to the black-swan theory (Taleb, 2010), and while we do have years of experience in IPv4 networks, we do not have the same amount of case studies for IPv6.

Literature does identify security as the "sum" of a number of properties like confidentiality, integrity, availability, authenticity and non-repudiation. We do prefer to use a different approach, as those are the properties that need to be ensured, but if we look at them alone we will probably end up missing big points. Any network (or any system) can be seen as the result of two main processes: design and implementation. This is valid for a whole network and its single components. In order to achieve the goal of enforcing the five above mentioned properties, a network has to be designed and implemented with specific features. The properties can then be enforced directly on the design and implementation or being built

on top of it. On the other hand, a flawed design (or implementation) will make all the efforts vain.

Another point to keep in mind is related to the concept of security itself. The term "security" is fancy and sounds good, but it is quite generic. According to the Oxford Dictionary, security is defined as *the state of being free from danger or threat.* The problem is then to identify what is the meaning of "danger" and "threat" and to relate with their prevention.

An aeronautical system is a Critical Infrastructure (CI) whose ultimate goal is to carry passengers and goods across the sky in a secure mode. A CI is, in general term, any infrastructure that is considered very important for economical or social relevance for a country. Any CI nowadays rely for its operations on a telecommunication system and on associated informatics systems. Those are called Critical Infrastructure Information (CII). It is worth noticing that the CII itself have little or no interest on the original goals of the CI. Instead, the CII is interested in meeting its own targets about security, performance, reliability and so on. The latters have to be driven by the CI mission, and this is the main problem in defining the security of a CII. Once those security requirements have been defined, the separation of concerns is completed and the CII can be defined as an almost completely orthogonal domain although being still a fundamental part of the CI.

To set the CII priorities in a correct and traceable way, it is of paramount importance to use an enterprise model usually in form of Enterprise Architecture built on top of an Enterprise Architecture Framework.

Even if the approach seems linear, in the particular case of aeronautical systems there are some peculiarities that have to be considered and, so far, are still open points:

- Each airport serves a number of different carriers,

- Each carrier lands in a number of different airports,

- Each carrier and airport is bound to follow:
    - International rules (aviation ones, mainly),
    - Local rules (national laws).

Each of the above should contribute to the definition of the EAF principles and drive the EAF constraints.

All these elements make it extremely difficult to find a common ground that might be useful and good for every single entity involved in the system.

In this chapter we will briefly describe the EAF principles and how they can be successfully used to model a CI. We will then outline what are the major differences between a "classical" IP network and an IPv6 network, and how these differences should be particularly be addressed in order to not pose a security risk for the CI.

2. Enterprise Architecture frameworks

The concept of enterprise can be applied to any type of organizations, commercial, public services, governments, and in general also to CI. The aims of an enterprise are to optimize all parts of the organization in a coherent way, rather than to achieve local optimization at business unit level (Sherwood et al., 2005). "The role of 'architecture' is to provide the framework that breaks down complexity into apparent simplicity" (Sherwood et al., 2005). EA is an abstract view which requires collaborative information from both business and IT professionals (Oda et al., 2009).

2.1 Importance of using an EAF in a Critical Infrastructures

Critical Infrastructures are large and complex enterprises, which often are composed of multiple systems. Organizations seek to establish common frameworks that allow them to rule on the development and evolution of such systems. The enterprise architecture frameworks are established templates for the development of enterprise architectures that aim to describe the enterprise structure, goals, processes and organization.

The CI protection process has the following steps:

• Identifying critical infrastructures essential for mission accomplishment.

• Determining the threats against those infrastructures by looking into business processes.

• Analyzing the vulnerabilities of the critical infrastructure.

• Assessing the risks of the degradation or loss of capability to achieve a mission.

• Defining countermeasures where risk is unacceptable.

• Defining security controls.

To help fulfill the steps it is necessary to identify critical business mission and then to conduct the classical steps of business process analysis, risk assessment and risk management.

2.2 The value of Enterprise Architecture framework

Enterprise Architectures (EA) are established approaches to developing descriptions of complex systems. An EA is often used to create a description of a structure and operations of an enterprise. An EA is a set of models that depicts how various business and technical elements work together as a whole. An Enterprise Architecture identifies the enterprise structure and gives a blueprint of its operation. The descriptions include multiple components, such as business drivers, principles, strategies, assets, technology and people. They also include selective views. Furthermore, an EA often addresses both current, future and interim states of an enterprise, including a transformation roadmap, change management process, program and portfolio context. EA describes the terminology, the composition of enterprise components, and their relationships with the external environment.

Enterprise Architecture needs a plan that defines a sequence of architecture states, also known as the transition architectures, that will change the existing ("as-is") EA to a desired target architecture. The EA roadmap is typically implemented through a number of projects, each delivering a solution architecture. These projects vary widely in scope and complexity.

Enterprise Architecture Frameworks (EAF) are common templates, and methods, for the development of instances of Enterprise Architecture (EA). They capture and represent all the relevant aspects of a business-driven enterprise. The known examples of such frameworks include the Department of Defence Architecture Framework (DODAF) used by the US DoD (*The DoDAF Architecture Framework Version 2.02*, n.d.), Federal Enterprise Architecture (FEA) used by the US federal agencies(*Federal Enterprise Architecture (FEA)*, n.d.). Zachman Framework (*Zachman Framework*, n.d.) and TOGAF (*The Open Group Architecture Framework 9 (TOGAF9)*, n.d.) can be considered enterprise architecture frameworks too. Figure 1 outlines the historical development of some major architecture frameworks.

2.3 Architecture Framework components

Typically, an Architecture Framework components include: views, meta-model, techniques, tools and guidelines for architecture development. It may also include best practices and design patterns. Other components can be used by architecture developers at their will.
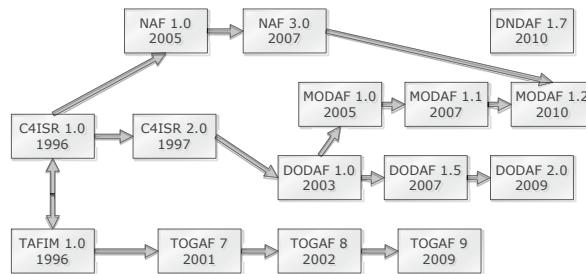
Fig. 1. Historical development of major Enterprise Architecture Frameworks
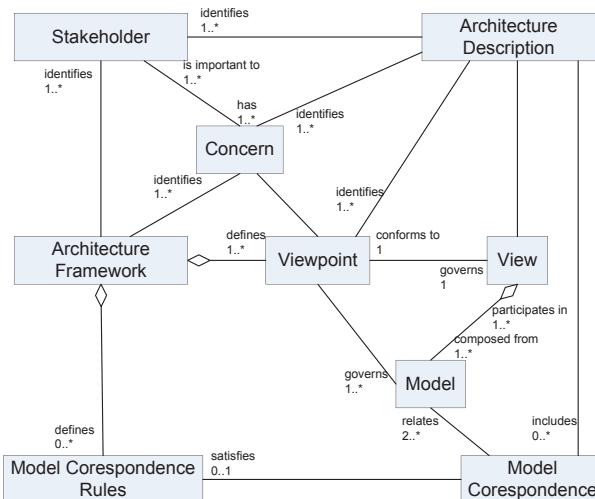
Fig. 2. Meta concepts used to describe architecture frameworks and their relationships

2.3.1 Viewpoints

Viewpoints are collections of useful views on the architecture data. They are selective distillations of complex architecture descriptions intended for the purpose of architecture presentation for different groups of architecture stakeholders. Viewpoints enable people to comprehend very complex systems. They limit the scope of presentations of various solutions to domain-specific issues, brought up by stakeholders involved in the architecture development. The views are concrete instances of the architecture data captured in a model. The aggregated architecture data that is represented in all the views from all the viewpoints could be considered a full composite model of the enterprise. This coherent set of architecture data is not always directly accessible as an artifact but is often scattered across different views. The model correspondence rules keep the composite model coherent and minimal.

2.3.2 Meta-model

Meta-model is a repository of terms that are used to represent the domain of interest. These are used during the AF development process to refer all the important aspects of the architecture. The meta-model serves, to a large extent, as common vocabulary. As such, it needs to be

complete, consistent, and minimal. Examples of meta-models are DM2 of DODAF2 and Content Framework from TOGAF.

A Meta-model is a critical element in an Enterprise Architecture Framework but it is even more important to also have a domain-specific Meta-model, which is understandable and widely accepted to all the stakeholders involved in the development and use of the Architecture Framework. Such an enhanced dictionary allows for a compact presentation of recommendations, procedures, methods and techniques used in a specific enterprise domain. There are some restrictions imposed on the integrity between the concepts. At various views (viewpoints) concepts are used to present some specific technique or procedure.

2.3.3 Techniques, guidelines, best practices and design patterns

The availability of viewpoints, consisting of multiple selective models and selective views of an enterprise, enables the incorporation of the domain specific knowledge into the framework. This is carried out with the use of the domain-specific concepts from the meta-model. During that activity, the set of models with predefined processes, enterprise design patterns (Wolthusen, 2004), and structures library are created. Security patterns are also reflected here. The typical scenarios such as security procedures are analysed and described with the use of framework concepts originating from meta-model and presented on the viewpoints defined in the framework.

2.3.4 The methodology and specific sub-methodologies

A generalized methodology is required for developing an enterprise architecture for critical infrastructures. It is required to produce such a description of a critical infrastructure architecture, so that different desctiptions be compatible and comparable.

2.3.5 Architecture framework's taxonomy

The architecture frameworks differ in type and scope. To understand better a given architecture framework, a hierarchical categorization is shown in Figure 3. The five levels position the domain-specific Architecture Framework between an organization-specific AF and the generic AF. Hierarchical layout inherits all the features of a generic Architecture Framework, but it also accommodates the methods that allow addressing systematically domain-specific concerns. 'Domain specific' means that terminology is typical of the particular domain interest - here: critical infrastructures. Typical processes and procedures found in CI and CII are identified, and the catalogue of good practices and guidelines is compiled together with common questions and their possible solutions.

There are five levels of AFs. The higher level the more abstract the AF is:

- Product Architecture (PA), often referred to as a solution Architecture, defines the architecture of a single product. It is usually developed for a single solution only and it is a natural way of product development.

- Organization-specific Architecture Framework (OAF) standardizes architecture description within a single enterprise, which allows it to have harmonic approach to development of multiple products and to maintain all the process in a uniform way.

- A Domain-Specific Architecture Framework (DAF) adds concepts which are common across similar organizations of a domain, and serves as a reference for OEM-specific architecture frameworks (e.g. PSAF). This approach specifically allows multiple organizations to cooperate easier together.
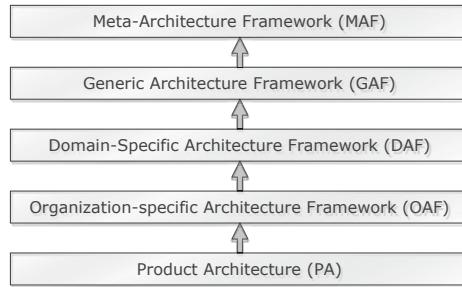
Fig. 3. Architecture Frameworks hierarchy

- The Generic Architecture Framework (GAF) defines concepts that are domain-independent (for example DODAF and TOGAF) and are universal enough to cover all the typical issues related to architecture development.

- The Meta-Architecture Framework (MAF) is independent of any type of system development. It provides a conceptual infrastructure to define and to reason about architecture frameworks. The major effort in this area is ISO standard 42010:2007 with extensions under development related to Architecture Frameworks.

The architecture framework for securing aeronautic communication can be considered as a domain specific architecture framework. It is not related with any particular product or organization, nor it is a general purpose framework or a meta-framework. The scope of this AF is the security in an aeronautical environment, and the purpose of building the architecture with this AF is to study and to reason on the protection of the communication infrastructure.

2.4 Existing Enterprise Architecture frameworks
2.4.1 Department of Defence Architecture Framework
The Department of Defense Architecture Framework (DoDAF) has been defined to serve as the structure for organizing architecture concepts, principles, assumptions, and terminology about operations and technology into meaningful patterns to satisfy specific DoD purposes. DoDAF version 2.0 focuses on architectural data, rather than on developing individual products as described in previous versions (DODAF 1.0, DODAF 1.5). The data-centric approach for architecture description allows for data re-use within and across different projects and over project life-cycles. It also supports flexible reporting and integration with other enterprise information systems. Finally strict representation of architecture data and their physical representation supports data sharing among multiple vendor tools for architecture development.

2.4.2 DoDAF data model
The architecture data in a described architecture is defined according to the DoDAF Meta-model (DM2) concepts, associations, and attributes. The terminology is used in similar way as defined in ISO 42010 and extensions regarding the architecture frameworks which are currently under development in ISO. The structure of architecture data is defined and constrained by DM2 data model. Figure 4 presents some selected example concepts from the DM2 Conceptual Data Model and their relationships. The concepts such as the Information, the Activity and the Performer are shown along with relationships between them.
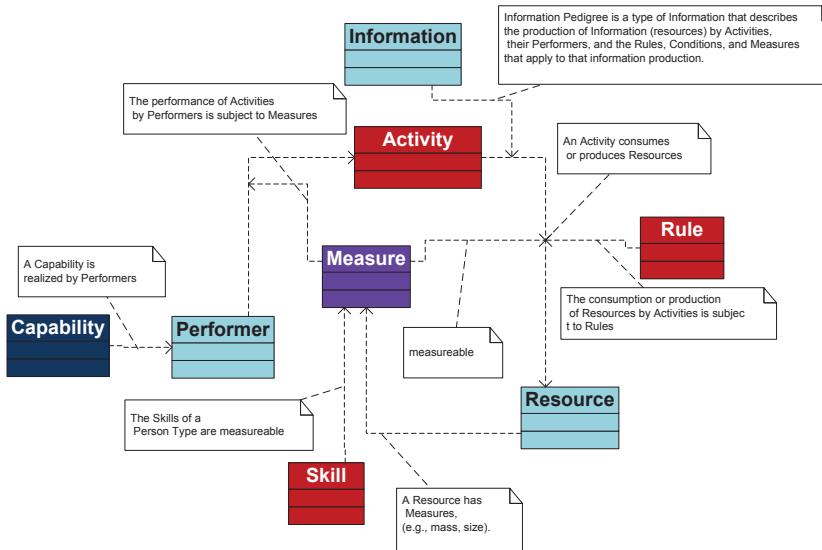
Fig. 4. Sample concepts and their relationships originating from DoDAF Data Model DM2

CDM consists of 26 concepts with their relationships.

2.4.3 DoDAF views, fit-to-purpose views

Views in DODAF v. 2.0 can be regarded as queries to the underlying architecture data. This is contrary to legacy versions of DODAF (1.5, 1.0) that were product (view) oriented. In DODAF2.0 the focal point is the underlying data - not the views. All relationships between objects are already contained in the architecture data, while views are only kind of queries. The core of DoDAF v. 2.0 is a data-centric approach where the creation of architectures to support decision-making is secondary to the collection, storage, and maintenance of data needed for efficient and effective decisions. The architect and stakeholders select views to ensure that architectures will explain current and future states of the process or activity under review. Selecting architectural views carefully ensures that the views adequately explain the requirements and proposed solution in ways that will enhance audience understanding. DODAF, similarly to MODAF and NAF, defines a collection of viewpoints:

- All Viewpoint (AV) - describes the overarching aspects of architecture context that relate to all viewpoints,

- Capability Viewpoint (CV) - articulates the capability requirements, the delivery timing, and the deployed capability,

- Data and Information Viewpoint (DIV) - shows data relationships and alignment structures in the architecture content,

- Operational Viewpoint (OV) - includes the operational scenarios, activities, and requirements that support capabilities,

- Project Viewpoint (PV) - describes the relationships between operational and capability requirements and the various projects being implemented,
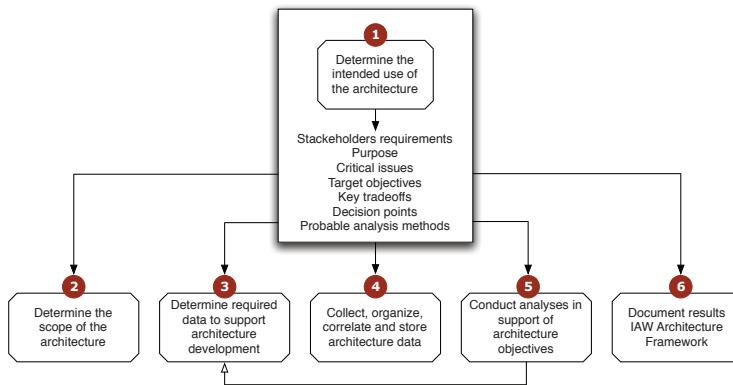
Fig. 5. DoDAF 2.0 Six-Step Process of Building an Architecture Description (Source: DoD Architecture Framework, Version 2.0, Volume 1, Section 2.1)

- Services Viewpoint (SvcV) - is the design for solutions articulating the Performers, Activities, Services, and their exchanges, providing for or supporting operational and capability functions,

- Standard Viewpoint (StdV) - addresses the policies, standards and sector recommendations across architecture,

- Systems Viewpoint (SV) - the design for solutions articulating the systems, their composition, interconnectivity, and context providing for or supporting operational and capability functions.

2.4.4 DoDAF methodology
DODAF 2.0 introduces a 6-step general methodology for architects to follow. The major steps are presented in Fig. 5.

2.4.5 The advantages of DODAF 2.0
- DM2 is based on the experience and the analysis of existing frameworks and methodologies. It is compliant with ISO/IEC 42010:2007.

- DODAF 2.0 is data-centric as contrary to product (artefact) centric. Accent is put on architecture data: methods for collecting, managing and exchanging all data related to created architecture. DM2 meta-model provides a common factor for all the products created while process. Consistent meta-model (DM2) is defined down to physical layer.

2.5 Generic EAFs - The Open Group Architecture Framework (TOGAF)
The Open Group Architecture Framework (TOGAF) is an example of a Generic Architecture Framework (GAF). It provides a user with a comprehensive approach to design, planning, implementation and governance of an enterprise architecture. TOGAF is developed and maintained by the Architecture Forum of The Open Group. It was originally developed in the mid-1990's, and has continuously evolved since then. The recent revision, TOGAF 9, can be downloaded from the web site of The Open Group. This section is descended from the

TOGAF specification and in many cases it's required to refer it for more detailed description of some particular issue.

TOGAF is a high level and holistic approach to design, which is typically modelled at four domains:

- Business Architecture - defines the business strategy, governance, organization, and key business processes.

- Data Architecture - describes the structure of an organization's logical and physical data assets and data management resources.

- Application Architecture - provides a blueprint for the individual application systems to be deployed, their interactions and their relationships to the core business processes of the organization.

- Technology Architecture - describes the logical software and hardware capabilities required to support the deployment of business, data and application services. This includes IT infrastructure, middleware, networks, communications, processing, standards, and so on.

TOGAF is comprised of six parts:

- Architecture Development Method (ADM) - an iterative sequence of steps to develop an enterprise-wide architecture.

- ADM Guidelines & Techniques - guidelines and techniques to support the application of the ADM.

- Architecture Content Framework - a detailed model of architectural work products, including deliverables, artifacts within deliverables and the Architecture Building Blocks (ABBs) that deliverables represent.

- Enterprise Continuum - a model for structuring a virtual repository and methods for classifying architecture and solution artifacts.

- Reference Models - the TOGAF Technical Reference Model and the Integrated Information Infrastructure Model.

- Architecture Capability Framework - a structured definition of the organizations, skills, roles and responsibilities to establish and operate an Enterprise Architecture.

2.5.1 TOGAF Architecture Development Method

The Architecture Development Method (ADM) is central to TOGAF. The ADM explains how to derive an organization-specific enterprise architecture that addresses business requirements. It includes establishing an architecture framework, developing architecture content, transitioning, and governing the realization of architectures. All of these activities are carried out within an iterative cycle of continuous architecture definition and realization that allows organizations to transform their enterprises in a controlled manner in response to business goals and opportunities. Structured as a series of nine phases plus requirements management phase that interacts with each of the nine phases throughout the ADM lifecycle, the ADM becomes an iterative method, over the whole process, between phases and within phases. Additionally, TOGAF, like most of the Architecture Frameworks, may introduce some new paradigms in the run of the architecture development process.

The phases within the ADM are:
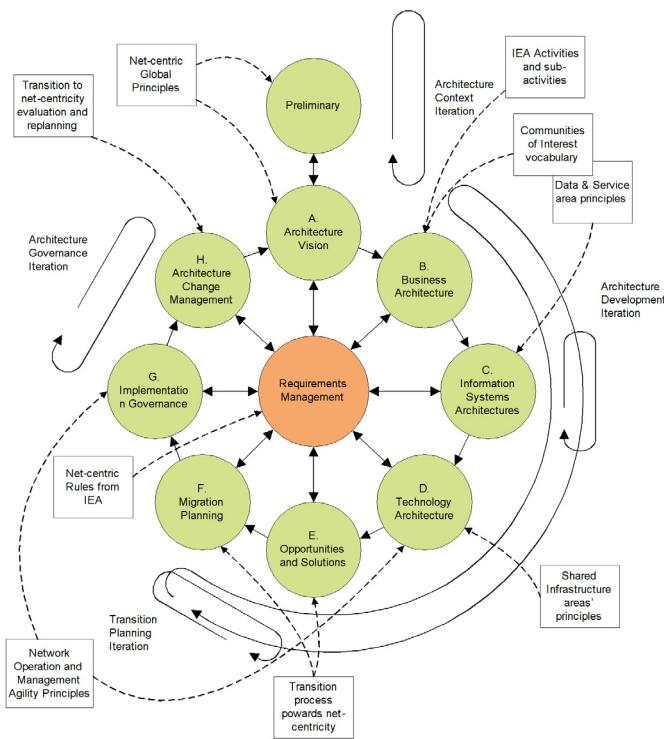
Fig. 6. The Architecture Development Method cycle

- The Preliminary Phase describes the preparation and initiation activities required to meet the business directive for a new enterprise architecture, including the definition of an Organization-Specific Architecture framework and the definition of principles.

- Phase A: Architecture Vision describes the initial phase of an architecture development cycle. It includes information about defining the scope, identifying the stakeholders, creating the Architecture Vision, and obtaining approvals.

- Phase B: Business Architecture describes the development of a Business Architecture to support an agreed Architecture Vision.

- Phase C: Information Systems Architectures describes the development of Information Systems Architectures for an architecture project, including the development of Data and Application Architectures.

- Phase D: Technology Architecture describes the development of the Technology Architecture for an architecture project.

- Phase E: Opportunities&Solutions conducts initial implementation planning and the identification of delivery vehicles for the architecture defined in the previous phases.

- Phase F: Migration Planning addresses the formulation of a set of detailed sequence of transition architectures with a supporting Implementation and Migration Plan.

- Phase G: Implementation Governance provides an architectural oversight of the implementation.

- Phase H: Architecture Change Management establishes procedures for managing change to the new architecture.

- Requirements Management examines the process of managing architecture requirements throughout the ADM.

There are four main iterations within ADM: Architecture Context iterations that allow initial mobilization of architecture activity by establishing the architecture approach, principles, scope, and vision. Architecture Definition iterations allow for the creation of architecture content by cycling through the Business, Information Systems, and Technology Architecture phases. These iterations also allow viability and feasibility tests to be carried out by looking at 'opportunities and migration planning'. Transition Planning iterations support the creation of formal change roadmaps for the defined architecture. Architecture Governance iterations support governance of change towards the defined Target Architecture.

2.6 How EAF reflects in security?

While Enterprise Architecture Frameworks and Enterprise Architectures are abundant - due to their wide use in enterprise governance - reports of their applicability in the domain of enterprise security are not common. In that particular domain there is currently a fast growing interest in information security management and information assurance. To this end, (i) the Sherwood Applied Business Security Architecture (SABSA) from SABSA Limited, (ii) Control Objectives for Information and related Technology (COBIT) from ISACA, and (iii) Information Technology Infrastructure Library (ITIL) from UK Office of Government and Commerce are currently the leading methodologies in information security and assurance, each focusing on similar problems, but emphasizing and addressing specific questions differently. Each strives to give detailed descriptions of a number of important practices and provides comprehensive checklists, tasks and procedures that an organization can tailor to its needs.

The Sherwood Applied Business Security Architecture (SABSA) is a model and a methodology for developing Enterprise Information Security Architectures and for designing security infrastructure solutions that support business needs (Sherwood et al., 2005).

SABSA relies on a model, which consists of a 6x6 SABSA matrix, where rows represent layers of the architecture model (contextual, conceptual, logical, physical, component and operational), while the columns represent the six major concerns in architecture modeling: assets, motivation, process, people, location and performance (Stango et al., 2011). These concerns are representations of Zachman's six communications questions, respectively: what, why, how, who, where, and when (*Zachman Framework*, n.d.). The topmost contextual layer allows to capture the context necessary to understand the requirements and the business attributes that shape the security required. The conceptual layer defines security concepts, principles and management procedures. In the logical layer, security controls are designed and the management procedures are specified. The logical security services are covered at the SABSA physical layer, in terms of physical security mechanisms. The component layer is related with the selection of products and technology. Finally, the operational layer is concerned with classical system operations work (Stango et al., 2011). The process of developing enterprise security architecture consists in populating the SABSA matrix by following SABSA workflows.

By comprising a number of models, frameworks, methods and high-level processes, SABSA allows to develop risk-driven enterprise information security and information assurance architecture. It can be used for the development of architectures at any level of granularity of scope. Being an open standard, it may be used in industry sector and in private and public organizations. To some extent SABSA is unique by being a risk-driven, enterprise information security and information assurance architecture. The layers of the architecture model and partitioning of concerns allows for two-way traceability of the architecture artifacts, thus allowing to check the architecture development product for completeness, to make sure that every business concern has been properly handled, and that the associated security requirements have been tackled with enough attention. SABSA also provides reverse traceability for business justification. It allows any architecture decision to be linked back to the original business requirements.

2.7 Common approaches to security

The IT industry has developed sets of standards which address security management.

2.7.1 ISO Standards

ISO/IEC 27000-series security standards are probably among the most commonly security standard deployed in enterprises worldwide. The series provides best practice recommendations on information security management, risks and controls within the context of an overall Information Security Management System (ISMS). The documentation structure and design are similar in design to management systems for quality assurance (the ISO 9000 series) and environmental protection (the ISO 14000 series) which were developed earlier. The current state of standard evolved from BS 7799 from 1995 published by BSI Group. The initial document was written by the United Kingdom Government's Department of Trade and Industry (DTI) and was composed of several parts. Currently, the set of ISO 27000 series standards include documents numbered 27000-27006, 27011, 27031, 27033-1 and more. Some selected ones are described below.

2.7.1.1 ISO 27002

ISO/IEC 27002 basically outlines controls and control mechanisms, which may be implemented subject to the guidance provided within ISO 27001 described below. The standard "established guidelines and general principles for initiating, implementing, maintaining, and improving information security management within an organization". The actual controls listed in the standard are intended to address the specific requirements identified via a formal risk assessment described with more details in ISO 27005. The standard is also intended to provide a guide for the development of "organizational security standards and effective security management practices and to help build confidence in inter-organizational activities".

2.7.1.2 ISO 27001

The significant reverse in numbering between BSI standards and ISO standards reflects the truth of the way the security awareness was developing. The increasing level of IT peril caused that simple in the beginning check-list were replaced with more complex and systematic approach. The new response to IT menace was the Part 2 of BS 7799 already mentioned which was adopted later in November 2005 by ISO and named ISO/IEC 27001.
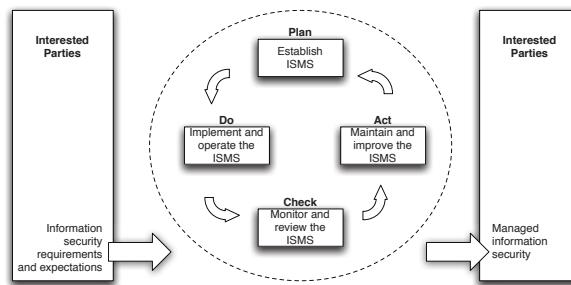
Fig. 7. Deming cycle applied to ISMS processes

The initial BS 7799 Part 2 (aka BS7799-2), entitled "Information Security Management Systems - Specification with guidance for use." was published by BSI in 1999. The main focus of BS 7799-2 was how to implement an Information security management system (ISMS). The document really brought a new quality to BS 7799-1 giving implementation guidelines to the information security management structure and controls identified in BS 7799-1. The most significant change of next revision form 2002 of BS 7799-2 is the Plan-Do-Check-Act (PDCA) (Deming quality assurance model), aligning it with quality standards such as ISO 9000. The phases or activities of the Deming cycle are:

- PLAN - Establishing the ISMS.
- DO - Operating the ISMS.
- CHECK - Monitoring and reviewing the ISMS.
- ACT - Improving the ISMS.

Even if periodical checks and proactive planning were present in some places of the previous versions BSI documents, its formal introduction brought significant change.
Other important aspects which are defined in ISO 27001 is the responsibility of the management for the ISMS, placing risk management as integral part or standard ISMS operation and economic justification of security-related expenses .

2.7.1.3 ISO 27005

The ISO/IEC 27005 standard, published in 2008, provides guidelines for information security risk management. It supports the general concepts specified in ISO/IEC 27001. It is designed to assist the implementation of information security based on a risk management approach. It does not specify, recommend or even name any specific risk analysis method. Its value is in specifying a structured, systematic and rigorous process from analysing risks to creating the risk treatment plan.

2.7.2 FISMA/NIST
The next important set of standards and best practices comes from USA. The Federal Information Security Management Act of 2002 called "FISMA" requires each federal agency to develop, document, and implement an agency-wide program to provide information security for the information and information systems that support the operations and assets of the agency, including those provided or managed by another agency, contractor, or other source.(*Federal Information Security Management Act (FISMA) Implementation Project*, n.d.)

FISMA similarly as described above ISO standards looks to the financial justification of the security means and explicitly emphasizes a "risk-based policy for cost-effective security." FISMA requires that agencies have an information systems inventory in place. The head of each agency shall develop and maintain an inventory of major information systems, including major national security systems, operated by or under the control of such agency. The guidance on determining system boundaries can be found in NIST SP 800-18, Rev. 1, Guide for Developing Security Plans for Federal Information Systems.

2.7.3 Categorize information and information systems according to risk level

According to FISMA, all information and information systems should be categorized according to the objectives of providing appropriate levels of information security according to a range of risk levels. The definitions of security categories are defined in the mandatory security standard FIPS PUB 199 "Standards for Security Categorization of Federal Information and Information Systems". The more detailed practical guidelines are provided by NIST SP 800-60 "Guide for Mapping Types of Information and Information Systems to Security Categories".

2.7.3.1 Security controls

FISMA requires that all federal information systems must meet the minimum security requirements, defined in the mandatory security standard FIPS 200 "Minimum Security Requirements for Federal Information and Information Systems". NIST Special Publication SP 800-53, "Recommended Security Controls for Federal Information Systems" defines the minimum security requirements which have to be met by organization by selecting the appropriate security controls and assurance requirements. The process of selecting the security controls to achieve adequate security is a multifaceted, risk-based activity involving management and operational personnel within the organization. Agencies have some choice in application the baseline security controls in accordance with the tailoring guidance. This allows agencies to adjust the security controls to more closely fit their mission requirements and operational environments. Practical implementation help guiding through the process can be found in NIST SP 800-53A "Guide for Assessing the Security Controls in Federal Information Systems and Organizations, Building Effective Security Assessment Plans". The controls selected or planned must be documented in the System Security Plan.

2.7.3.2 System security plan

Agencies are obliged to develop policy on the system security planning process. NIST SP 800-18 introduces the concept of a System Security Plan - a collection of living documents that require periodic review, modification, and plans of action and milestones for implementing security controls. Procedures should be in place outlining who reviews the plans, keeps the plan current, and follows up on planned security controls.

Without having the System Security Plan a proper security certification and accreditation process for the system is impossible.

2.7.3.3 Risk assessment

FIPS 200 along with NIST SP 800-53 requires a foundational level of security for all federal information and information systems. The agency's risk assessment validates the security control set and determines if any additional controls are needed to protect agency operations.
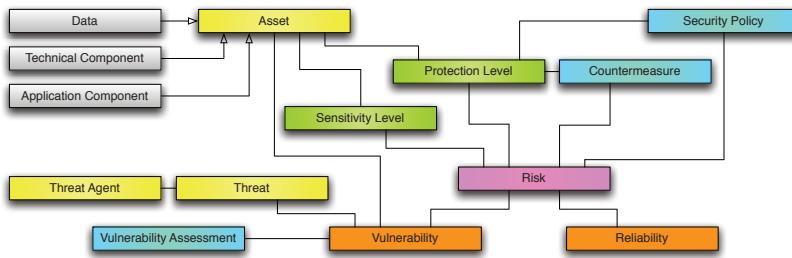
Fig. 8. Risk Analysis Model

2.7.3.4 Certification and accreditation

The certification and accreditation process is defined in NIST SP 800-37 "Guide for the Security Certification and Accreditation of Federal Information Systems". Security accreditation is the official management decision given by a senior agency official to authorize operation of an information system. The set of mandatory conditions which system must fulfill in order to receive it includes: proper system documentation, completed risk assessment and the review of system's controls to be functioning appropriately.

2.7.3.5 Security standards comparison

If we try to compare the approaches to security of the ISO 27k series and NIST Special Publications dealing with security we can see that they are influenced by each other. They are close to each other in their motivation and objectives, but differ in the primary focus, and in the level of details. The target reader is not the same, too, which causes some important changes in their use. The NIST standards are mandatory for all federal agencies by law so they are limiting choice of an agency where ISO series mentions only possibilities as addressed to a business not bound by legal regulations. Other difference is caused by the budgeting. The level of details and structural organization of the NIST series results in noticeably higher level of details, and quality. Also these documents reflect, to a larger than the ISO series, the new approaches and methods.

3. Risk Assessment

EA can be useful to describe the structure of the Enterprise and to map it to architectural and technical components able to fulfill the Enterprise goals. The security of the whole Information architecture, however, needs to be analyzed in a little more specific detail.

When analyzing the security of the CII, the model in Figure 8 should be used. In this model the prime component is the definition of Assets, as is the union of the Data, Technical (hardware and software) components and the Application Components. The difference between the latter two is the use of Data: whereas the Applications modify and use Data assets, the Technical components does not.

In this model every asset have some desired properties (i.e., its Protection and Sensitivity levels) and some inner properties (i.e., its Reliability and Vulnerability). The Sensitivity Level is defined by the importance of the asset for the operations of the CI, while the Protection level is defined and is the outcome of the Risk Analysis process. The sum of Countermeasures and Security Policies contribute to quantify this property. Thus, it is not an intrinsic but rather an

extrinsic property. The most interesting part, probably is the Vulnerability, as is the asset's resistance to faults due to attacks or misbehaving entities.

In the diagram are shown also the Security Policies and the Countermeasures. Both are actively contributing to define the Risk by lowering it to acceptable levels. The main differences between the two are that while the Security Policies are aimed at preventing a dangerous event, the Countermeasures have the target to react to an event and mitigate the consequences.

Furthermore, assets can be divided into *primary* and *secondary*, depending on their relative importance for the CI operations. According to the EA results, a set of scales based on the threat estimation and the vulnerability of each asset have to be set, and opportune policies (Security and Countermeasures) should be applied in order to lower the risks to acceptable levels (also to be defined in the EA).

A fundamental part of this process is thus the evaluation of the Threats and the Vulnerabilities.

3.1 Threat estimation

The Threat estimation is a very difficult part, as it involves assumptions on the Threats like the exploitability of a Vulnerability, the capabilities of an attacker and so on. It is extremely important to have a realistic Threat Model (Rescorla & Korver, 2003), otherwise the threat estimation might lead to over or underestimation. An extremely important point, however, is to never consider exploitation probability as dependent on the knowledge of the vulnerability itself. Always assume that the attacker have the maximum knowledge possible.

3.2 Vulnerability Assessment

Vulnerability Assessment (Thompson & Chase, 2005) is an integral part of the Risk Analysis process. Each Application or Technical Assets should pass some tests in order to measure their functionalities. The most common ones are the conformance tests, while vulnerability analysis tests are less used.

The conformance testing aims at verifying that the system is behaving correctly to a set of known inputs. It is a common practice to require a conformance against some protocol or some reference implementation in order to ensure interoperability. Vulnerability testing, on the contrary, aims at discover unexpected behaviours when the system have unexpected inputs. On a simpler scale one can consider a vulnerability test as a search for backdoors, possible bugs and so on. Since this kind of tests involve unknown variables, they are usually much more complex and costly than the conformance tests. Nevertheless it is imperative to adopt some sort of vulnerability assessment system, as those are exactly the kind of vulnerabilities an attacker could use to violate a system. At the moment the vulnerability testing is by far the most difficult and challenging part of an asset verification. Ni2S3 EU project (http://ni2s3.kt.agh.edu.pl) developed a methodology and a full framework for Vulnerability Assessment. The interested reader can check Ni2S3 outcomes and deliverables.

4. Vulnerabilities in IPv6 networks

Although EAF can (and should) be used to describe the structure of the Information systems and the network, when it comes to deploying a lot of problems might occur. Whilst most of them are "known", we think that IPv6 might be one of the major risks, mainly due to underestimation. In this section the main vulnerabilities in IPv6 networks are pointed out.

4.1 Differences between IPv4 and IPv6

We will assume that the reader is familiar enough with IPv6, so we will not describe IPv6 details. The goal here is to summarize the main points that are related to security.

The main and, probably, most known difference between IPv4 and IPv6 is the addressing space size. Although this is one of the key points, it is not at all the most important one. Among the new IPv6 features, there are some more interesting and under evaluated features that might be a concern for security.

In the pure spirit of the Internet of Things concept, IPv6 designers decided to push the auto-configuration features to the limit, allowing a near-seamless plug and play network model. This has been reached by defining and enforcing the concept of auto-configuration for all the relevant network layer features, from IP addresses acquisition to network knowledge (e.g., routers, gateway, DNS discovery).

4.1.1 IP addresses

The structure of the IPv6 address is described in (Hinden & Deering, 2006). An IPv6 address is logically divided into two main parts: a network part and a node address part. Each of them is (or can be) auto-configured.

Multicast and Anycast addresses are particularly important, as they play a major role in the security of the IP protocol. As one could expect, a multicast address is a one-to-many address. The relevant point is about the scope of this address. In IPv6 multicast can be restricted to having a local or global scope (or more complex scopes, not to be described here).

About anycast addresses, it is interesting to consider the definition (Hinden & Deering, 2006):
An IPv6 anycast address is an address that is assigned to more than one interface (typically belonging to different nodes), with the property that a packet sent to an anycast address is routed to the "nearest" interface having that address, according to the routing protocols' measure of distance.

Multicast in IPv4 is a quite rarely used system. On the other hand in IPv6 it is used to contact routers, DNS, address configuration and so on. Anycast is a completely new addressing scheme in IPv6. Their importance is related to their use in network operations, as all the plug and play IPv6 features rely on them. Due to this, it is quite evident how a malicious user could leverage them to attack the network infrastructure and cause potential damages.

4.1.2 IP address acquisition

IPv4 defines various methods to get a valid network address. However, if two network elements try to use the same IP address, there is no automatic way to fix the conflict. In IPv6 the changes are radical about this point. First and foremost, the main thing to keep in mind is that a single interface has always more than one address, and they are all valid.

Any networked entity has at least one link-local IP address and one multicast address per interface. The first is algorithmically built and allows communications across the link (either a point-to-point or a switched network), the second is closely related to the first and is used to solve the possible conflicts between nodes that, by chance, might end up with the same address.

To clarify this, it is worth explaining how an address is built. There are a number of ways to build a valid address, but the most used are: 1) Auto-configuration, 2) DHCP, 3) Manual configuration. The first of them is probably the most used. It provides a number of ways to build a valid address. The first and probably the easiest way is to use the MAC address as part of the IPv6 address. This, however, have the drawback of identifying almost uniquely the

device, so its communications can be tracked by a malicious user in a quite easy way. Another approach is to randomly choose the address (MS Windows does that), however in this case the drawback is the total opposite: you can not identify the entity by its address anymore, and every time there is a network restart, the address will be different.

Another interesting way is to use the so-called Cryptographically Generated Address (CGA) (Aura, 2005; Bagnulo & Arkko, 2006), where part of the address is a hash of a public key. Although the use of CGA is very interesting and can be used in a number of ways, their processing does require an extra load in the routers. Hence, they can be successfully exploited to trigger fancy network attacks.

In any case, auto-configuration does require a protocol to solve possible conflicts among the addresses, and this is handled by the Neighbor Discovery protocols (Arkko et al., 2005; Narten et al., 2007). The discovery is based heavily on multicast, and a malicious user can use this as well in order to trigger a denial of service attack (see section 4.3.1).

Thus, in the auto-configuration case, each node is able to build a valid link-local address in the form of FE:80::[auto-configured 64 bits address]. In order to gain a global-scope address, i.e., an address valid for the global Internet, the entity shall contact a router through a well-known multicast group and receive a valid prefix. Again, this is a nifty but potentially dangerous feature.

DHCP approach is not different from the IPv4 one, but it can also be used by autoconfigured nodes to acquire the DNS address. DHCP as well can be a vulnerability, as it relies on a well-known multicast address.

4.1.3 IPv6 address scope

The last key point to keep in mind when dealing with IPv6 networks is the absence of Network Address Translators (NATs). Although NAT was originally proposed as a technique to delay the inevitable IPv4 address exhaustion, it quickly became a way to separate network segments and "hide" parts of it from the public internet. NATs, however, can be bypassed in a number of ways, some actually raising quite dangerous exploit possibilities (Huston, 2004). In IPv6 there is no need of NATs (the address space is large enough to use global addresses), even tough for some security or multihoming configurations a NAT could be still a viable solution (see (Thaler et al., 2010)).

The real point, however, is that all the IPv6 hosts can have a *global scope* address, thus exposing them to the traffic from the public internet. As a rule of thumb, the network planners/administrators should keep in mind that everywhere there was a NAT in IPv4, in IPv6 there should be a firewall. Moreover due to the global address use, it becomes imperative to implement Intrusion Protection (or Detection) Systems, so to quickly react to unexpected user's behaviours. It should be expected, as a matter of fact, an increased number of peer-to-peer systems and connections, not easily blocked by firewalls, and not at all unwanted per-se.

4.2 Vulnerabilities of the IPv6 infrastructure

In order to reduce the security to a manageable problem it's useful to divide it among its basic components The basic level of separation shall be between the vulnerabilities arising from the design and the ones related to the implementation.
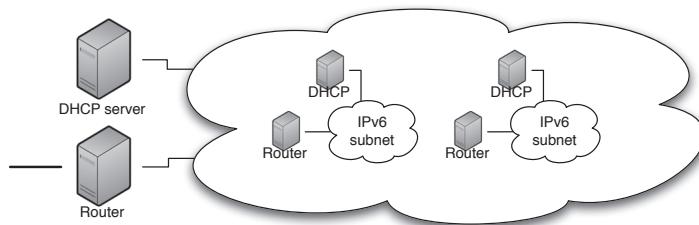
Fig. 9. Subnetwork hierarchy

4.2.1 Design flaws

The first and probably the most important aspect is about the network design. There is not a single way to build an IPv6 network and each design might lead to potential security issues.

From a general point of view, there is no real difference between an IPv6 and an IPv4 network at LAN level. The main difference arises when we look at the larger "network", depicted in Figure 9. As we might notice, there is a hierarchy of routers and DHCPs servers. Also this architecture might not seem different from a classical IPv4 network, however in IPv4 most of the subnet routers would have been NATs and the local DHCPs administratively disjoint from the main DHCP server.

Due to the lack (or disappearance) of NAT, the role of network segmentation goes to routing and firewalling. On the other hand, the lack of NATs makes it central the role of firewalls. These have to be placed practically in every single router, and their configuration has to be consistent.

On the other hand, only a handful of firewall/routing configuration managers are able to properly handle IPv6 routing tables, and this can be quite a problem.

A second point is about the address configuration policies. Some network parts could need fixed addresses, while some other might need a more flexible auto-configuration mode. Due to the differences in address configuration, the firewall rules might be quite complex to setup.

Again about network obfuscation, some policies for the DNS have to be adopted. It is a well-defined policy in IPv4 to have reverse-address available by default, so that the DNS knows about all the possible network addresses. This is clearly impossible for IPv6 networks due to the address space. This poses a design problem. Should the DNS store all the numbers or just the relevant ones? It is out of the scope to give an answer, but it is reasonable to assume that the DNS should store the direct and reverse mapping only for the well-known addresses, as is the manually configured and fixed ones. The other ones should be left unmapped. It is of course possible to update the DNS in real-time coupling it with the DHCP (if any), but this does not seems to give any real benefit beside keeping this "fake" address existence.

Another design point is about address assignment delegation. In large networks it seems reasonable to divide the network in sub-networks, much like in IPv4. The existence of global scope address is based on the Router Advertisement process (i.e., a router will periodically or on-demand provide RA messages with the valid network prefix). A frequent design flaw is thus related to the decision about administratively separated domains, as is about the subnetwork topology. It is rather obvious that each network part should not have more

than one router answering with RAs.[1] Network design should carefully choose the size and topology of the subnets in order to avoid excessive signaling overhead for each subnetwork. A wrong design might expose the network to Denial of Service attacks or unexpected failures. About the DHCPs, it is rather obvious that their parameters should be consistent. IPv6 do admit the use of DHCP delegates, or slaves. In this case, as in the previous one, the signaling overhead should be minimized and checked.

Last but not least, there is the issue related to where and how to insert security probes in the network. While the previous design decisions where mainly related with the overhead analysis, as is "intrinsic" faults due to overheads, the probes should look for anomalies in the network. Thus, it is necessary to look for the main possible architecture attacks.

4.2.2 Architectural attacks

The main attacks possible aimed at disrupting the IPv6 architecture use "creatively" the plug and play IPv6 features. Unfortunately any decentralized or consensus system is somewhat subject to attacks, and IPv6 is not different.

- Router's advertisements can be forged, and a malicious user (or a malfunctioning unit) could inject in the network false or wrong RAs. The possible attacks range from Denial of Service to Man-in-the-Middle.

- DHCP architecture suffers the very same issue. A malicious user could impersonate the local DHCP and inject false data in the network. According to the address construction methods this could lead to different attack kinds.

- The last attack is related to the ND protocol (Narten et al., 2007). IPv6 nodes must, periodically, check for duplicates in the network. A malicious node can easily exploit this mechanism to implement a Denial of Service simply replying to all Duplicate Address Detection messages.

Those three kinds of attacks make it clear the importance of continuously monitoring the network in order to promptly discover attackers or malfunctioning nodes. It is also rather important to point out that also simply misconfigurations or malfunctioning nodes could exhibit the above behaviours. The network administrator shall not assume that the absence of attackers make the network immune from those issues. More insights on the specific attacks will be in section 4.3.

4.2.3 Implementation flaws

The kind of attacks toward IPv6 implementations are mostly arising from bugs or vulnerabilities in the IP or in the application stacks, and should, in theory, not be too difficult to find and eliminate. On the other hand there is a bad habit among the implementors, as is to give for granted that the underlying software is working as intended. If this might be true for IPv4 stacks, it is not anymore so for IPv6. As a matter of fact, IPv4 stacks have been in use for decades, so there are well-known and well-tested implementations that seldom are changed. On the contrary IPv6 has been around for decades as well, but with much less testing and use. Moreover, the additional functionalities in IPv6 like autoconfiguration, multicast, anycast, IPSec, etc. make the protocol quite more complex. Hence, it can be expected that some implementations might contain bugs or non-optimized code (the latter can lead to DoS).

---

[1] A special case is where there are double or triple exit points, or fallback routers, but this case is rather specific.
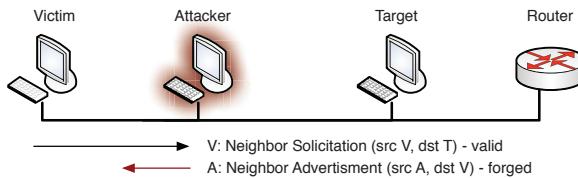
Fig. 10. Neighbor Discovery attacks

At L4 and above layers, as is TCP, UDP, Application Level Protocols, etc., things are not looking brighter. Theoretically IP stacks should be agnostic with respect to the IP version, with application level protocols "thinking" in terms of URIs and not bothering with actual IP numbers. In practice there are a number of cases where this is not possible or simply the programmers did violate this principle. A good rule of thumb is to never consider a properly working in IPv4 system (i.e., passing vulnerability and conformance testing under IPv4) as "safe" for IPv6. Tests should be re-evaluated and considered as independent.

In order to minimize the impact of implementation flaws, two main techniques should be considered: conformance testing and vulnerability testing.

4.3 IPv6 specific attacks

This section will summarize some new attacks specific to IPv6. The aim is not to be exhaustive, but to show what kind of threats can be expected reasonably in an IPv6 infrastructure. Moreover the attacks listed here are well known, so one can expect that an attacker will try them.

4.3.1 Address resolution attacks

In IPv6 the ARP protocl is substituted by Neighbor Discovery (ND) protocol (Narten et al., 2007). To resolve the IP - MAC address mapping, two new packets exists: ICMP6 Neighbor Solicitation / Neighbor Advertisement. The use is quite straightforward: if A needs B's MAC address it will send an NS using multicast "solicited-node" address (or unicast, depending on the link capabilities). B will reply with an unicast NA. Since anyone can reply to the NS message, an attacker can forge NA replies and claim to be either the victim or even all the machines in the network. The countermeasure is quite obvious, but it requires to monitor the network continuously. Moreover false positive could arise frequently, as it is quite difficult to find a generic rule to spot this kind of attack.

Another "flavor" of this attack is related to the IP assignment procedure. Any host, at boot time, will initiate a double Duplicate Address Discovery (DAD) procedure, in order to make sure no duplicate address is in the local network. This is particularly important in case of auto-assigned addresses, but it is used also in case of DHCP assigned addresses. The attacker can send an ICMP NS and pretend to be any host in the network. In this case the result is a Denial of Service to the victim, as it will not be able to acquire any IP address. The basic ND attacks scheme is depicted in Figure 10.

It should be noted that, since NA can be also be unsolicited, in case an interface changes its IP address, this kind of attack can be also target working machines, triggering unnecessary DAD procedures. If used on a whole network it can be quite dangerous.
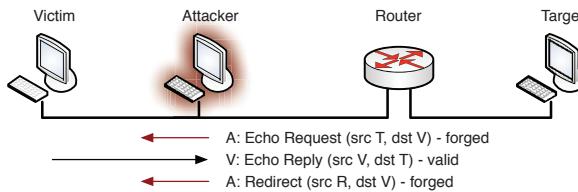
Fig. 11. ICMP Redirect attack

4.3.2 Router attacks

The attacks involving routers are strictly related to the previous set of attacks. In this case the attacker can use ICMP Router Advertisement in order to claim to be a router, or to inject in the network false prefixes. The result is different, as in the first case the attacker will become a router, thus allowing an easy man-in-the-middle. In the second case the network will be blocked. A different approach is to implant a bad route through creative use of ICMP redirect. This kind of attack is a bit trickier than the previous ones, but it is still quite simple. Figure 11 shows this attack.

A similar class of attacks involves the use of DHCP. As DHCP are in IPv6 replying to requests on a specific multicast address, also in this case an attacker can forge DHCP replies in order to inject in the network fake informations. In particular the DNS association is sensible, as one could point to an almost-valid DNS, meaning a DNS replying correctly to all the requests but some specific ones, thus redirecting only some specific addresses to a fake server.

This kind of attacks are quite dangerous and are well documented (see (Nikander et al., 2004)), but they have something in common: they all come from local network. There are some passive countermeasures, like SEND (SEcure ND, (Arkko et al., 2005)), but SEND use can not be considered as a definitive solution, as its applicability is not universal.

The best countermeasure to those attacks is for sure to secure the local network, not allowing an attacker to join the network. This can and should be done by disabling the physical unused ports and monitoring the network for unexpected behavior of the local hosts. An implementor should never consider the local network as secure, and always check for compromised hosts. An attacker gaining control of an host in the local network can easily block it.

4.3.3 Traffic amplification attacks

This class of attacks can be triggered either from local network or from remote network. They all involve triggering automatic replies to legitimate messages, like Echo Requests, to the specified victim, flooding its interfaces. Another target could be a specific link, in an attempt to consume all the available bandwidth. A quite handy way to do that involves using Routing headers in order to force a communication between two controlled hosts (even external to the attacked network) passing through one of the attacked routers. The two hosts can generate enough traffic to fill the available link bandwidth.

4.3.4 Fragmentation, mobile and tunnel attacks

One misconception in IPv6 is about fragmentation not existing anymore. It is true that fragmentation has been changed, but it is still possible to have fragmented packets.

In IPv6 the *routers* are not allowed anymore to fragment, but the source can still use it. Hence, fragmentation and reassembly is limited to the source and destination hosts and it is used

when the L4 layers does not (or can not) honor the discovered MTU size. An attacker can use this in order to mask a malicious routing header, hence it is imperative to defragment the packets in a firewall in order to inspect the real packet content and block dangerous communications.

About Mobile IPv6 (see (Johnson et al., 2004)) attacks, the protocol itself is to be considered secure, as it involves a massive use of IPSec in order to protect the communications. On the other hand all implementations have an option to disable IPSec. The implementor should carefully check to reject non secure communications, as any unprotected Mobile IPv6 link can be easily redirected to a different destination.

Last but not least, tunnel attacks do involve injecting packets in a IPv6 over IPv4 link. Also in this case the tunnel should be protected with proper cyphering, otherwise an attacker can inject packets easily just guessing the two tunnel endpoints.

4.3.5 Dual stack attacks

Attacks involving dual stack hosts (i.e., hosts with both IPv4 and IPv6) are quite common, but not so different from the other ones presented before. They have been separated mainly because they do involve network design and management in order to be coped with.

The point is that IPv6 and IPv4 infrastructure could be different due to NAT use in IPv4. Hence, the firewalls could be placed in different points or have different setups. This of course should be avoided as much as possible, but the problem of keeping firewall rules consistent between the two domains still remains. The network administrator should always keep in mind that the attacker could use a double stack attack (i.e., part of the attack on IPv4 and part on IPv6) in order to gain access to a victim host. Hence, any security solution should consider as a priority to keep consistent rules in both domains.

On a different perspective, it should be kept in mind that most O.S. have dual stacks already implemented and IPv6 ready to run. The Administrators should disable the unwanted stack or, at least, make sure that the wanted one have precedence. On this point, it is worth pointing out that IPv6 is normally the preferred stack by default. This last point could not be the case for aeronautical networks, where IPv6 use is mandatory, but it could be still a problem for part of the network in the airport. As a general rule IPv4 traffic should not be allowed outside limited areas (e.g., public internet areas). Hosts should not be dual stack enabled unless necessary and appropriate synchronized security policies should be used.

4.3.6 Network discovery attacks

The last consideration concerning peculiarities about IPv6 is the network discovery procedure. In IPv4 an attacker had a number of ways to discover the network structure, the most known being network scanning through a brute-force search for assigned IP numbers. This was feasible in IPv4 since the number of hosts in a subnet is typically quite limited, so it was possible to use pings or port scanning on all the IP numbers of a LAN.

In IPv6 this is simply not feasible, as the number of hosts to scan for is typically extremely large. A "normal" IPv6 subnet is a /64, meaning the attacker would had to scan about 2^{64} hosts (more than 18 millions of millions). This means that a brute-force discovery would take around 500 millions of years. Using clever techniques one could lower the time to some months, but it is still clearly not a feasible attack. This point, however, should not give a sense of false security, as the attacker will probably try to recover the needed informations from public-available sources: the DNS.

The network administrators and security planners should always keep in mind that the attackers will target:

- Public hosts, discovered through google, DNS, etc. and Anycast address hosts.

- All standard services hosts (DHCP, routers, time servers, etc.) through local multicast addresses.

- All hosts though local multicast (can be still time consuming but is feasible).

The most interesting point is about the DNS. It is expected that DNS servers will be one of the major source of informations, hence the public DNS should never contain any detail about the "internal" hosts. Even tough internal hosts could have names and globally valid IP addresses, their presence should not be made public unless really necessary. Moreover whatever does not need a name resolution but only an IP address should not, from a security perspective, have a canonical name. This will increase network obfuscation and will make it more difficult for the attacker to find the network structure.

5. AAA architectures

AAA stands for Authentication, Authorization and Accounting and is an important area in commercial telecommunication networks. There are, however, some misconceptions about AAA that should be considered, as its application is not consistent in all networks, nor it is equally considered important. Especially in the Internet community AAA systems are seldom considered as an important part of the network.

First we should point out the differences between each 'A'. *Accounting* is generally confused with billing, thus omitted in the networks where there is no need for it. Nevertheless, Accounting is mainly responsible for resource usage tracking, which can be used for a number of other purposes like network planning and resource optimization. The second set of mistakes is around the confusion between Authentication and Authorization. Keeping things simple, *Authentication* is the process of validating an entity identity. *Authorization*, on the other hand, is about the verification of the entity rights to use a specified resource. As an example, consider to have to use a network printer. The printer could ask an Authentication (e.g., ID and password), then it could check through Authorization if you can use all its features or just a subset (e.g., print in color or just black and white) and finally it could log the resource usage (number of pages printed) through Accounting. The very same model can (and should) be applied on network use, differentiating access privileges (e.g., the rights to use all Internet ports, firewall setups and so on).

IEEE 802.1x (see Figure 12) is one of the most widely used AAA frameworks, at least in the Internet. We can identify three actors:

- the *Supplicant*: the entity needing an Authentication.

- the *Authenticator*: the entity the Supplicant is authenticating with.

- the *Authentication Server*: the entity actually performing the Authentication.

It is worth noticing that the Authenticator and the Authentication Server are different functions, and must be separate hosts for security purposes. As a matter of fact the Authenticator needs the Supplicant to be authenticated, but does not require to know anything about the Supplicant for real. All it needs to know is if the Supplicant has been authenticated and if it has the rights to access a particular resource.
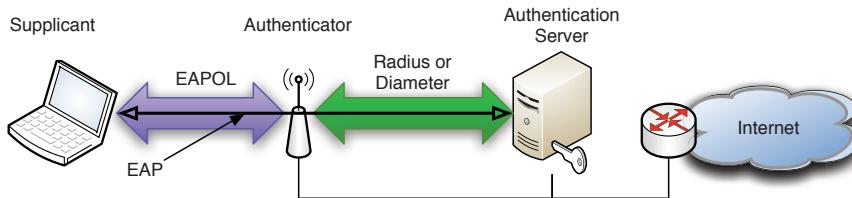
Fig. 12. 802.1x schematic architecture

The process is rather simple. The Supplicant needs to send a set of credentials to the Authenticator. The credentials are contained in an EAP (Aboba et al., 2008) envelope. EAP messages are exchanged between the Supplicant and Authenticator through the EAPOL protocol. The Authenticator will forward these to the Authentication Server using either RADIUS (Congdon et al., 2003; Rigney et al., 2000) or DIAMETER (Calhoun et al., 2003) protocols. The Authentication Server will send back to the Authenticator the keys needed to start a proper secure channel. EAPOL is not limited to the IEEE 802.11, but it can be used in IEEE 802.3 (Ethernet) and other kinds of media. The Authenticator can be any entity requiring an authentication / authorization from the user in order to use its resources.

The architecture, per-se, is not so complex. On the other hand, from a security perspective, a number of aspect must be take into account. It is fairly clear that the Authenticator does not need to know any information about the Supplicant identity, and should not either. On the other hand the Authenticator will have access to the EAP messages exchanged between the Supplicant and the Authentication Server. Hence, it is imperative that these messages are secure, meaning a compromised Authenticator must not be able to discover the Supplicant authentication informations. This is particularly important, as EAP itself defines a number of methods (about 40) to identify the Supplicant and/or the Authentication Server. Some of these methods are less reliable than the others, allowing man-in-the-middle or password discovery though rainbow tables (e.g., MSCHAP). Unfortunately there is not a clear and simple solution about which EAP method to use, as the various devices can have different supported methods. The only suggestion in this case is to make sure to disable the unwanted methods (i.e., the ones considered too weak) in the Authentication Server, and to prefer methods based on challenges and/or hardware, like EAP-AKA.

The second point to be taken into account is about the Authentication Server itself. It seems worth noticing that this network element is for real the most important one to be secured. Since it is the core of all the authentication and authorization process, no other services should run on the host in order to minimize the risk of a successful attack on the server. Moreover its architecture should be redundant and failsafe.

AAA protocol requirements have been defined in (Aboba et al., 2000). The basic requirements are: (1) Scalability, (2) Fail-over, (3) Mutual authentication between the client and the server, (4) Transmission level security, (5) Data object Confidentiality, (6) Data object Integrity, (7) Certificate transport, (8) Reliable AAA transport mechanism, (9) Run Over IPv4 and IPv6, (10) Auditability, (11) Ability to carry service-specific attributes. About this, it is necessary to outline the two architecture of RADIUS and DIAMETER, as they are quite different, although similar.

5.1 RADIUS protocol and architecture

The RADIUS protocol was created in 1997 with Dial-In users as primary target. The protocol run over UDP and defines a set of messages to authenticate the user: *Access-Request*, *Access-Accept*, *Access-Reject* and *Access-Challenge*. The Access-Challenge is used to request additional information to the Client in order to complete the authentication. The authentication information are included in a list of Attribute-Value pairs. The protocol itself defines some of them, but there is the capability to define vendor specific ones. On the other hand the Attribute is coded in a single byte, so only 255 attributes are possible.

RADIUS defines also an Accounting mechanism (Rigney, 2000), moreover it is possible to define roaming policies by using *Realms*. A Realm is defined by prepending or appending a Realm information to the user identity, e.g., `userid@company.com` or `company.com\userid`. Upon receiving a request, a RADIUS Proxy will compare the realm information with a table and will forward the request to an appropriate RADIUS server. It is worth noticing that the realm is a simple test string and does not need to pair with any actual real internet domain.

From a security point of view, RADIUS shows some weakness. First and foremost, the cyphering method used in messages is quite weak (MD5), so all the connections must be tunneled via stronger methods, like IPSec. The second point is about authentication. There is no mutual authentication between the client and the server. Hence, the client must "trust" the server. While this might as well be solved through secure tunnels, it is still a weakness. The third point is concerning UDP. Since the transport layer is unreliable, complex timeouts are necessary, and there is no guarantee that a request is being processed. Last but not least, RADIUS does not supply any "easy" method to provide failover or redundancy. Hence, it have scalability issues.

Despite the above mentioned points, RADIUS is one of the most used protocols for AAA in Internet, with worldwide AAA federations (e.g., eduroam (*Eduroam - Education Roaming*, 2011)) serving users all around the world. Moreover, it is a well-supported system, with a good user-base and many implementations, thus providing some reliability, at least for what concerns the limitations and implementation / deployment knowledge.

5.2 DIAMETER protocol and architecture

DIAMETER has been developed in 1998 to overcome the limitations of RADIUS. The main difference is the use of TCP and SCTP instead of UDP. DIAMETER has been chosen by 3rd Generation Partnership Project (3GPP) as the AAA protocol for IP Multimedia Subsystem (IMS) (3GPP, 2011), and should be expected to be the reference protocol in future AAA IP systems.

DIAMETER supports application-layer acknowledgements, and defines failover algorithms and the associated state machine. It makes IPSec mandatory, thus enforcing transport-level security, and defines explicit roles for Agents (Proxies, Redirects and Relays). Moreover it defines a correct framework to support Capability negotiation and Auditability, among a number of other features. Last but not least, the attributes space has been increased in order to support a greater number of vendor-specific data.

Despite the obvious superiority of DIAMETER, its use in Internet is still a niche, with IMS implementations being the only real test case. This is probably due to the lack of support by clients and the lack of free and supported servers, with only a handful of exceptions. Nevertheless, from a general point of view, it is fairly clear that the RADIUS weaknesses

should suggest to work toward its substitution, so any implementor should consider RADIUS only as a temporary solution.

6. Conclusions

Aeronautic communication can be treated as a kind of critical infrastructure and as an enterprise. In order to enable the protection of this infrastructure it is mandatory to identify the enterprise mission, the organization structure, critical business processes and relationships inside and outside. The more complete picture of the enterprise the better, because the context and interdependencies need to be properly reflected in enterprise description. The description, will have an impact on the infrastructure security planning and deployment. There exist known means to prepare enterprise descriptions - these are enterprise architectures. The security planning and deployment should be also based on a deep knowledge of the threats, the threats model and, most important, of the possible vulnerabilities an attacker can exploit. As integral part of the risk analysis process, vulnerability assessment tools should always be used, especially for IPv6 architectures. Last but not least, the implementors should always take particular care about the new IPv6 features, as the most common risk is to underestimate the changes in IPv4 to IPv6 transition.

7. Acknowledgments

8. References

3GPP (2011). TS 23.228 - IP Multimedia Subsystem (IMS); Stage 2.

Aboba, B., Calhoun, P., Glass, S., Hiller, T., McCann, P., Shiino, H., Zorn, G., Dommety, G., C.Perkin, B.Pati, D.Mitto, S.Mannin, M.Beadle, P.Wals, X.Che, S.Sivalingham, A.Hamee, M.Munso, S.Jacob, B.Li, B.Hirschman, R.Hsu, Y.Xu, E.Campell, S.Baba & E.Jaques (2000). Criteria for Evaluating AAA Protocols for Network Access, RFC 2989.

Aboba, B., Simon, D. & Eronen, P. (2008). Extensible Authentication Protocol (EAP) Key Management Framework, RFC 5247.

Arkko, J., Kempf, J., Zill, B. & Nikander, P. (2005). SEcure Neighbor Discovery (SEND), RFC 3971.

Aura, T. (2005). Cryptographically Generated Addresses (CGA), RFC 3972.

Bagnulo, M. & Arkko, J. (2006). Cryptographically Generated Addresses (CGA) Extension Field Format, RFC 4581.

Calhoun, P., Loughney, J., Guttman, E., Zorn, G. & Arkko, J. (2003). Diameter Base Protocol, RFC 3588.

Congdon, P., Aboba, B., Smith, A., Zorn, G. & Roese, J. (2003). IEEE 802.1X Remote Authentication Dial In User Service (RADIUS) Usage Guidelines, RFC 3580.

Eduroam - Education Roaming (2011).
URL: *http://www.eduroam.org/*

Federal Enterprise Architecture (FEA) (n.d.).
URL: *http://www.whitehouse.gov/omb/e-gov/fea/*

Federal Information Security Management Act (FISMA) Implementation Project (n.d.).
URL: *http://csrc.nist.gov/groups/SMA/fisma/index.html*

Hinden, R. & Deering, S. (2006). IP Version 6 Addressing Architecture, RFC 4291.

Huston, G. (2004). Anatomy: A look inside network address translators, *The Internet Protocol Journal* 7(3).

Johnson, D., Perkins, C. & Arkko, J. (2004). Mobility Support in IPv6, RFC 3775.

Narten, T., Nordmark, E., Simpson, W. & Soliman, H. (2007). Neighbor Discovery for IP version 6 (IPv6), RFC 4861.

Nikander, P., Kempf, J. & Nordmark, E. (2004). IPv6 Neighbor Discovery (ND) Trust Models and Threats, RFC 3756.

Oda, S., Fu, H. & Zhu, Y. (2009). Enterprise information security architecture a review of frameworks, methodology, and case studies, *Computer Science and Information Technology, 2009. ICCSIT 2009. 2nd IEEE International Conference on*, pp. 333 –337.

Rescorla, E. & Korver, B. (2003). Guidelines for Writing RFC Text on Security Considerations, RFC 3552.

Rigney, C. (2000). RADIUS Accounting, RFC 2866.

Rigney, C., Willens, S., Rubens, A. & Simpson, W. (2000). Remote Authentication Dial In User Service (RADIUS), RFC 2865.

Sherwood, J., Clark, A. & Lynas, D. (2005). *Enterprise security architecture: A Business Driven Approach*, CMP Books.

Stango, A., Pacyna, P., Rapacz, N. & Prasad, N. (2011). Proposed Risk prioritization based on SABSA (Sherwood Applied Business Security Architecture) for Critical Information Infrastructures, *3rd International ICST Conference on Security and Privacy in Mobile Information and Communication Systems*.

Taleb, N. N. (2010). *The Black Swan: The Impact of the Highly Improbable*, 2nd edn, Random House.

Thaler, D., Zhang, L. & Lebovitz, G. (2010). IAB Thoughts on IPv6 Network Address Translation, RFC 5902.

The DoDAF Architecture Framework Version 2.02 (n.d.).
URL: *http://cio-nii.defense.gov/sites/dodaf20*

The Open Group Architecture Framework 9 (TOGAF9) (n.d.).
URL: *http://www.opengroup.org/togaf/*

Thompson, H. H. & Chase, S. G. (2005). *The Software Vulnerability Guide*, Charles River Media.

Wolthusen, S. (2004). Modeling critical infrastructure requirements, *Information Assurance Workshop, 2004. Proceedings from the Fifth Annual IEEE SMC*, pp. 101 – 108.

Zachman Framework (n.d.).
URL: *http://www.eacoe.org/*

Quality of Service Management and Interoperability

Christian Kissling and Tomaso de Cola
German Aerospace Center (DLR)
Oberpfaffenhofen,
Germany

1. Introduction

Within this chapter quality of service management strategies are assessed with respect to their applicability and efficiency in the ATM context. In particular addressing the service demands of ATM communication, such as strict latency and loss limitations is considered herein. This also covers the architecture for selection of links for data transmission and the interaction between technology independent and technology dependent components in the networking architecture by means of standardized communication protocols such as IEEE 802.21 and ETSI BSM extensions.

The term Quality of Service (QoS) is used in a variety of different ways and often depends also on the context that it is used in. One notion of QoS denotes the performance of a service from the users view. A measure for the grade of QoS is how good the performance attributes of a service match with the demands made on it. The kind of attributes which are relevant and need to be fulfilled depends thus naturally on the context of the service. While for many other services perceived or qualitative QoS measures are applicable, the ATM communication environment envisaged here makes high and precise demands on different attributes, presented later on in detail. For common internet applications such as HTTP, email and VoIP a lot of work has been spent to map the quality perceived by the user to networking parameters which can be measured and controlled. (ITU-T, G.1010) provides for instance a model for multimedia QoS categories from the end user perspective. Tables with technical requirements are provided for eight different application categories such as audio, video and data services. In the literature a large number of publications are available dealing with different aspects of providing QoS in the terrestrial Internet but also in satellite networks. (Marchese, 2007) provides a good overview and summary of different aspects of QoS provision in the context of heterogeneous networks, such as present in the ATM environment considered here.

The provision of QoS in an operational and safety critical aeronautical environment is however considerably different from the applications and demands in the Internet. Service parameters such as defined in (ITU-T, G.1010) thus cannot be directly applied here. The most intuitive reason for this is that a violation of QoS attributes in Internet applications results in a reduced service quality, which is naturally undesirable and bothersome for users, but has not necessarily implications on operational events and safety of life. In the aeronautical domain, for the management of air traffic this is decisively different. Late

arrival of e.g. directive commands issued by the controller for the pilot can have catastrophic effects. Also corrupted messages or multiple receptions of messages can have such serious consequences, affecting the safety of the airplane and the passengers. For this reason it is not sufficient if the QoS mechanisms for ATM communication try to achieve the requirements as far as possible but it is necessary that the requirements are definitely met.

1.1 Network design

Since IPv6 is the unification point in the SANDRA network (SANDRA, 2011), there is the need of the design and adaptation to an aeronautical internet. Main focus within this task is the handling of the network management and also of the resource management. Additionally, effort is spent for the development of new and efficient handover and mobility management algorithms and concepts, respectively. Also an IPv6 based naming and addressing architecture will be provided. Due to the high degree of mobility on a global scale and the heterogeneous network environment (i.e. short-range and long-range terrestrial as well as satellite access technologies), work on a network mobility (NEMO) based IPv6 protocol started in contrast to the ICAO chosen Mobile IPv6 protocol supporting only host mobility.

For the SANDRA Terminal as shown in the aircraft segment of Fig. 1, the lower layer (data link and physical layers) functions are provided by an on-board Integrated Modular Radio (IMR) consisting of heterogeneous radio access technologies.
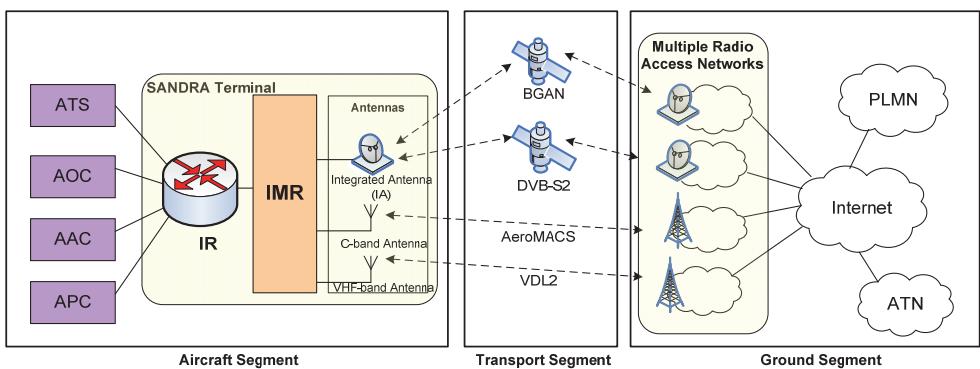
Fig. 1. SANDRA Network Architecture (Ali, 2011).

The upper layer (layer 3 and above) functions are managed by an Integrated Router (IR). The following chapter will describe in detail the realization of the connection of these two entities.

2. Quality of service management and interoperability

2.1 QoS definition for aeronautical networks

In a joint study of EUROCONTROL and the Federal Aviation Administration (FAA), potential future communication technologies which are suitable to provide the necessary safety and regularity of flight have been investigated and requirements for the future application services have been derived. The results of this study have been published in the so called "Communications Operating Concept and Requirements for the Future Radio

System (COCR)" (EUROCONTROL, 2007). Within this study the concepts of ATM have been analyzed from an operational point of view and the expected technical requirements have been formulated, also for services which are not yet deployed but are expected to be deployed in the future. The results in the COCR provide information for all operational services with respect to their periodicity, volume and technical requirements. The main QoS requirements for the services are the following ones:

- **Transmission delay (TD$_{95}$):** The TD$_{95}$ represents the one-way latency requirement for every Operational (OP) message which 95% of all messages of a service have to arrive within. It is defined per flight domain (i.e. Airport (APT), Terminal Maneuvering Area (TMA), En-Route (ENR) and Oceanic, Remote and Polar (ORP)), per service type (ATS and AOC) and for each Class of Service (CoS).
- **Expiration Time (ET):** In case the TD$_{95}$ is not met due to various reasons (e.g. packet loss) the COCR sets a so called Expiration Time within which the packets have to arrive which failed the TD$_{95}$ requirement. To be compliant with the requirements, the percentage of messages indicated by the continuity requirement has to arrive within the ET.
- **Continuity:** Denotes the probability that a transaction will be completed having met specified performance. With respect to the ET, this probability represents the percentage of the transmitted messages which arrive within the latency performance requirement set by the ET.
- **Integrity:** Denotes the acceptable rate of transactions that are completed with an undetected error. This requirement refers to packets which are considered to be received correctly but actually contain false information, e.g. caused by undetected bit errors
- **Availability:** Denotes the probability that the equipment comprising the system is operational and conforms to specifications (excluding planned outages and logistics delays). It is further distinguished into
  - *Availability of use*: Probability that the communication system between the two parties is in service when it is needed.
  - *Availability of provision*: Probability that communication with all aircraft in the area is in service.

The COCR specifies these QoS requirements per service, but also for aggregated Classes of Service (CoS). For the definition and evaluation of the QoS architecture, the three main impacting requirements are thus the **TD$_{95}$**, the **ET** and the **Continuity** requirement. Table 1 shows an excerpt from the COCR, specifying the ET, TD$_{95\text{-FRS}}$ and Continuity (C$_{\text{UIT-FRS}}$) for the different defined CoS.

Within the COCR, the different application services are then also mapped to the CoS listed in Table 1.

It should be noted that these requirements are impacted by the QoS architecture, but not entirely defined by it. Primarily the requirements are dependent on the underlying link technology which set boundaries for latency, packet loss etc. with the available data rate, propagation delay, retransmission mechanisms and forward error correction (FEC) methods. Clearly a QoS architecture cannot ensure compliance with the requirements if the underlying link and physical layer are not capable to transport the data sufficiently. On the other hand, in case the underlying link technology is providing sufficient transmission capabilities, the QoS architecture has to ensure that these abilities are efficiently used so the requirements are met. One challenge of the SANDRA design is thus to define a QoS

architecture which allows meeting the requirements, provided that the underlying link technology provides sufficient performance (in terms of throughput, latency and packet loss rate).

CoS	ET [s]	$TD_{95\text{-}FRS}$ [s]	$C_{UIT\text{-}FRS}$	Service Type
DG-A	Reserved	9.8	Reserved	NET Data
DG-B	1.6	0.74	0.99999992	ATS A/G Data
DG-C	5.0	1.4	0.9996	
DG-D	7.8	2.4	0.9996	
DG-E	8.0	3.8	0.996	
DG-F	12.0	4.7	0.996	
DG-G	24.0	9.2	0.996	
DG-H	32.0	13.6	0.996	
DG-I	57.6	26.5	0.996	
DG-J	Not available	13.6	Not available	AOC A/G Data
DG-K		26.5		
DG-L		51.7		

Table 1. Excerpt from the COCR CoS definitions.
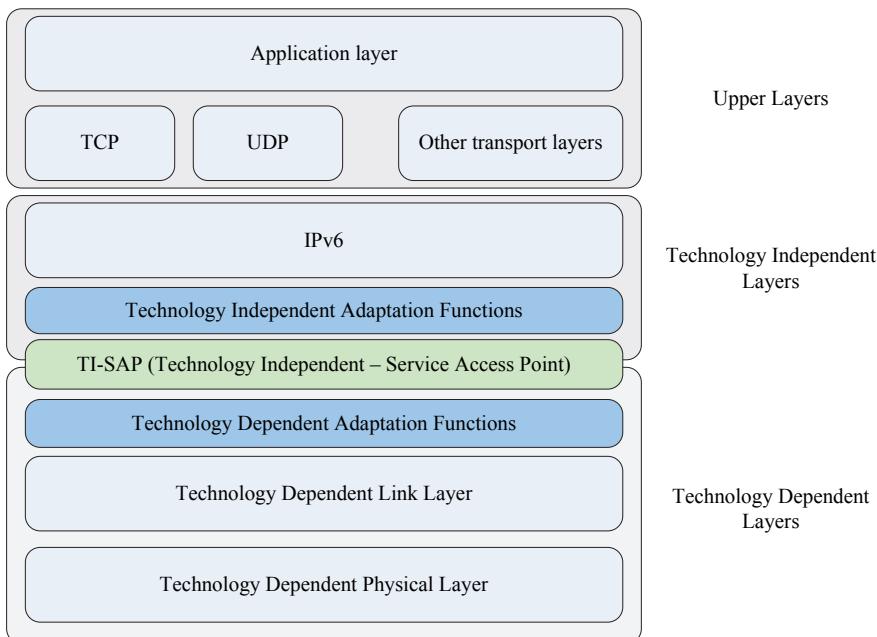
Fig. 2. Functional interaction between technology independent higher layers and technology dependent lower layers.

For the SANDRA QoS design, the additional problem is addressed how different communication links can be integrated into a seamless network and which mechanisms and approaches are suitable to allow provision of the required QoS. SANDRA hereby focuses on the network layer QoS mechanisms mainly. Fig. 2 illustrates the general approach. One requirement for the layer 3 QoS mechanisms is that they must be interoperable and independent of the type of used link. Going beyond this, also the uniform interfaces (denoted Service Access Points, SAP in the following) to the technology dependent Layer 2 are in the scope of SANDRA and discussed hereafter in more detail.

2.2 QoS mapping in the SANDRA architecture

As straightforward from the considerations drawn in the previous section, the necessity for the SANDRA architecture is to simultaneously manage different QoS traffic profiles and transmission technologies over which different services have to be handled, translate into a QoS mapping problem. Beside the technical challenges that arise in selecting the Layer 2 queues to which the traffic has to be forwarded depending on the QoS requirements (scheduling and QoS mapping problem), a particular attention has to be reserved to the characteristics of the QoS architecture, being embedded in SANDRA. Apart from the specific QoS model being adopted (IntServ or DiffServ as sketched in the following sections), some attention has to be addressed to how Layer 3 and Layer 2 intercommunicate, by preserving the QoS requirements specified in the Service Level Specifications (SLS) of the specific traffic service. In this respect, different approaches can be applied. Ad-hoc solutions can be deployed, by extending for instance the functionalities and the related primitives already available from the ISO/OSI protocol stack. Given the scope of the SANDRA framework, it is instead better to have a model in line with architectures currently or going to be standardised. In this perspective, the features offered by the ETSI BSM protocol architecture are worth being considered. The main peculiarity consists in the definition of the SI-SAP interface, virtually separating the upper layer (Satellite Independent, SI) from the lower layers (Satellite Dependent, SD) and providing dedicated primitives to efficiently manage QoS, Address Resolution and Multicast functionalities over satellite. The overall ETSI BSM protocol architecture is depicted in Fig. 3, where the main components are:

- SI layer: it implements the upper layer and in particular the IP protocol (versions 4 or 6). It also incorporates the Satellite Independent Adaptation Function (SIAF) module, which is responsible for adapting the SI functions to the characteristics of the lower layer specification, through dedicated primitives.
- SD layer: it implements the lower layer, in particular the datalink and the physical ones. It also implements the Satellite Dependent Adaptation Functions (SDAF) module, which interacts with the aforementioned SIAF through dedicated primitives.
- SI-SAP interface: it logically separates the SI from the SD layers, providing a set of dedicated primitives, exchanged between the SIAF and SDAF modules, responsible for QoS, address resolution and multicast functionalities.

In this light, it is reasonable to extend the principles of the ETSI BSM protocol architecture for application in the SANDRA framework, to particularly address the QoS requirements of aeronautical networks (Plass,2011).

In fact, two main "ingredients" of the SI-SAP interface can be re-used and properly extended to match the requirements of the SANDRA functional architecture: the Queue Identifier (QID) and the QoS primitives. The former is defined in the ETSI BSM protocol architecture as identifier of the Layer 2 physical queues, so to allow an efficient QoS mapping between Layer

3 and Layer 2 queues, through the dedicated QoS primitives. The latter, in turn, allows actually implementing the QoS mapping algorithms and offering the essential tool to perform the resource allocation, based on the requests coming from the upper layers. The QoS problem in the SANDRA network involves not only resource allocation issues but also transmission technology selection, thus requiring the extension of the current SI-SAP interface functionalities along with the use of the IEEE 802.21 architecture in terms of the Media Independent Handover (MIH) functions. In practice, the QID has to be conceptually extended in a way that it incorporates both queue and link identifiers. Besides, the integration and the interaction of the ETSI BSM and the IEEE 802.21 architecture is of primary importance to perform the communication of the link selection to the upper layer and perform the resource allocation based on the requirements notified from the higher layers (e.g., application protocol or management plane). To this end, the SI-C-QUEUE primitives will be conveniently extended in their scope so to also include the new functionalities, thus allowing the different components to interwork properly according to the SANDRA network characteristics.
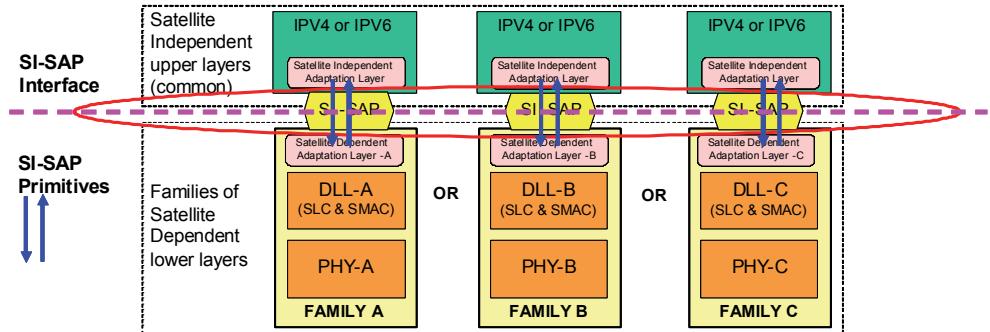

Fig. 3. ETSI BSM protocol architecture and SI-SAP interface definition.

At this point, the final point to be addressed is the way the described protocol architecture integration (ETSI BSM and IEEE 802.21 namely) can be finally embedded in the real architecture of the SANDRA network. In this respect, a particular attention has to be reserved to the IR and IMR interaction. Although the SI-SAP interface has been conceived to logically separate the upper from the lower layers within a satellite terminal, it can be easily extended to physically separate two different components, by distributing the implementation of the primitives. This can be done by re-thinking the SI-SAP interface as separating IR and IMR; these, in turn, will implement the related QoS primitives, thus working as the SIAF and SDAF modules in the original ETSI BSM architecture.

The overall system function can be then summarised in the following operations:

- In case the QoS requirements are constrained to a specific link by the upper layer, the IR will signal the selected transmission technology along with QoS request in a dedicated QID to the IMR, which in turn will forward the forthcoming data traffic to the specified transmission link. The availability of the transmission link is known after the start-up phase, which is accomplished by suitably combining the SI-C-QUEUE-open primitives with the MIH functionalities.

- In case no link-constrained request is performed by the upper layer, the IR simply signals the IMR about the QoS requests. In turn, the IMR will be responsible for running the link selection algorithm to identify the transmission technology most appropriate to match the received QoS requests. Also in this case the signalling is performed through

real exchange of the SI-SAP primitives; in particular, in this case the QID will basically contain an identifier for the QoS request and a default value of the transmission technology, being it not explicitly selected by the upper layers.

- In case a link was no longer available or its availability was reduced (upon notification through the specific MIH functions), the IMR would in turn notify it to the IR through the corresponding enhanced SI-C-QUEUE primitives to trigger a new resource allocation. The IR in turn will run a new resource allocation request to match the new link configuration, by modifying or demanding the assignment of a new QID.

The overall interaction between the SANDRA components is represented in the following picture.
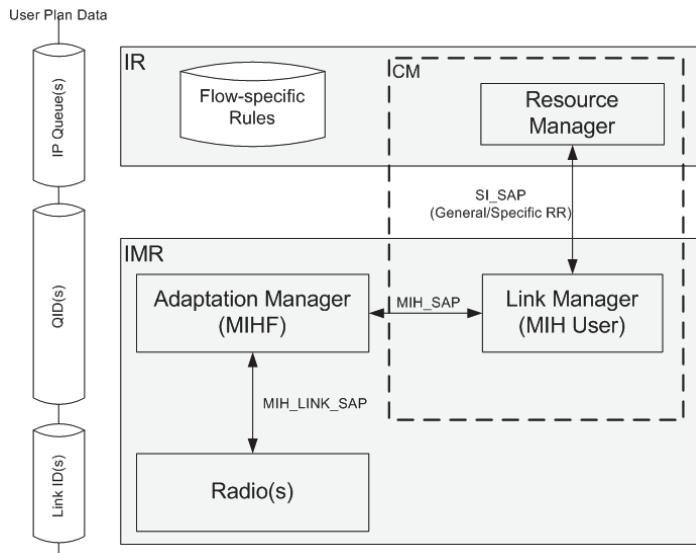

Fig. 4. Interaction between IR and IMR modules within the SANDRA network.

A particular attention has to be reserved to the interaction between IR and IMR in terms of message exchange, performed through primitives' generation and reception according to the architecture above described. In more detail, as it was introduced in the previous paragraphs, the overall IR-IMR system behaviour can be regarded as a sort of Master-Slave interaction, where either the IR or the IMR play the role of master and slave respectively, depending on the specific case being dealt with. In case the application is requesting specific link and QoS profiles, the IR plays the role of master, implying that the IMR will attempt to match the IR requests in terms of link allocation and resource management. On the other hand, when the link selection is forced by the IMR (which plays the master in this situation), the IR is basically responsible for forwarding data through the appropriate logical interfaces to the IMR, without taking any decisions in terms of data filtering and QoS policing/shaping.

The overall interaction can be described as a block diagram, where the two entities (IR and IMR) take decisions based on their role and the functions they are implementing.

The block diagram is essentially composed of the IR and IMR state machines:

- IR: It does not perform any operations unless the request of new radio resource is either issued by the IMR or by other external entities, such as application requests.

- IMR: It does not perform any operation unless a link is available (link label) or the allocated resources need to be updated.

Starting from these two states for IR and IMR, respectively, it is possible to exemplify the dynamics of the overall SANDRA system in presence of constrained and unconstrained services.

As far as the former is concerned, the IR will specify a new radio resource with a specified radio technology. This will be then notified to the IMR through proper primitives, which will be responsible for checking the availability of the requested resources as part of the radio resource management operations. In case the resource are not available, a loop of message exchange between IMR and IR is then initiated to agree on a different resource request, thus possibly ending up with the data forwarding operations.

As far as the latter is concerned, the radio resource request issues without specifying any radio technology, which will be instead selected by the IMR. Accordingly, the IMR is then in the position to setup the selected radio technology and performs the bandwidth allocation upon resource availability, following the same procedure reported before.

An additional case, independent of the specific service being handled, worth being considered is imminent handover or available resource change event. They are both handled by the IMR, which informs the IR through the appropriate primitives. In turn, the IR will update the radio resource assigned to a given service, by issuing a new request to the IMR; in order to match the current resource availability. It could be also the case that when no agreement is reached about the resources that a service should require (e.g., no alternative links are available), the service could be not admitted (or dropped) in (from) the SANDRA network.

The block diagram is sketched in the following picture, Fig. 5.
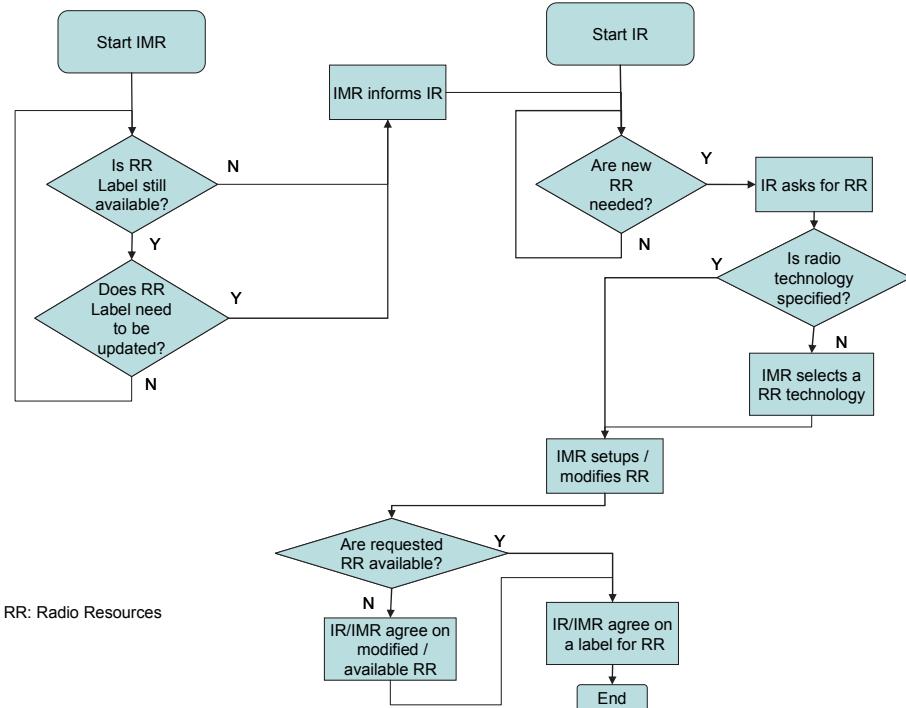
Fig. 5. IR/IMR interaction.

2.3 QoS management architecture

In contrast to QoS architectures which are deployed in the internet, the QoS design in the aeronautical scenario has to comply with a range of security and safety requirements which limit the freedom of choice for a QoS architecture considerably. The selected QoS management architecture should also rely on well established and standardized solutions. From today's perspective, one of the major design constraints is the strict separation of operational (ATS and AOC) and non-operational (AAC and APC) services within the network due to safety. While this separation is a real requirement nowadays, in SANDRA an all-integrated, seamless network is envisaged for the far future, which integrates also operational (OP) and non-operational (NON-OP) services and provides the required safety at the same time. Naturally this has also an impact on the QoS architecture.

Within the SANDRA context the challenge of integrating different communication links into a single common network architecture creates the need to deploy adequate QoS management functionalities. The QoS disciplines which have to be considered in particular for such a QoS architecture design include the following:

- **Connection Admission Control (CAC):** Technique used to decide which traffic is admitted into the network. Going back to the Asynchronuous Transfer Mode it is defined as "the set of actions taken by the network during the call set-up phase (or during call re-negotiation phase) in order to determine whether a connection request can be accepted or should be rejected (or whether a request for re-allocation can be accomodated)" (Hitoshi, 1998). In the SANDRA context, the rejection of an OP connection request is clearly not an option. In the scenario where OP and NON-OP domains are fully separated CAC is thus not applicable. When looking into the fully integrated scenario however, CAC is a technique which can be applied to the NON-OP domain to control the amount of traffic admitted from NON-OP sources that is injected into the overall network with the purpose to avoid disadvantageous impact on the OP services. The notion of "connection" can hereby refer to different aspects, e.g. to acceptance/rejection of users entering the system, of TCP connections, SIP connections or general data flows. The use of CAC techniques is supposed to increase the QoS perceived by the users since, e.g. the interruption of a voice call is perceived worse than a rejection of the call in the first place. For these reasons the application of CAC techniques should here be limited to NON-OP traffic.

- **Congestion Control (CC):** In case too many packets are present in a network the performance in terms of delay and loss rate (e.g. due to buffer losses) degrades. This situation is commonly called congestion (Tannenbaum, 2002). While for moderate levels of traffic load (i.e. injected packets) the packet delivery increases proportionally with the load, at some point the message processing is no longer able to cope with the packets, queue sizes first increase (at the cost of increased delay) and finally packets are dropped due to buffer overflow. When retransmission mechanisms without control are present, the packet drop will result in an even higher offered traffic load which in turn results in more dropped packets. Congestion control defines techniques which have the purpose of controlling the occurrence of congestion and ensuring that the network is able to carry the offered traffic. In contrast to flow control techniques, congestion control is a global issue involving all involved nodes. Within SANDRA, the network must be able to cope with the traffic offered by the OP services in any case. In other words the network must be sufficiently dimensioned so congestion due to OP traffic cannot occur. For a scenario where OP/ NON-OP networks are separated, CC is thus

not applicable for the OP part. In the all-integrated scenario, however CC is applicable to the NON-OP traffic in order to ensure that the NON-OP traffic cannot cause congestion in the network which is adversely affecting the OP services.

- **Scheduling / Queuing:** Packets are stored in queues until they can be transmitted on the link. The scheduling algorithm then determines the order in which the queues are served, i.e. the order according to which packets of the different queues are sent. For DiffServ architectures in the Internet, commonly known queues are Expedited Forwarding (EF), Assured Forwarding (AF) and Best Effort (BE). Many different scheduling algorithms are known from the literature, such as First In First Out (FIFO), Round Robin (RR), Weighted Round Robin (WRR), Weighted Fair Queuing (WFQ), Deficit Round Robin (DRR), Stochastic Fair Queuing (SFQ) or Worst Case Weighted Fair Queuing (WF²Q). For wireless communication, other scheduling algorithms taking the wireless nature of the medium into account are defined, such as for instance Idealized Wireless Fair Queuing (IWFQ), Wireless Packet Scheduling (WPS), Channel Condition Independent Fair Queuing (CIF-Q), Wireless Fair Service Algorithms (WFS) or Proportional Fair Queuing (PF). These lists are clearly not exhaustive but shall provide only a fundamental overview. Within the aeronautical scenario, the queuing and scheduling has especial importance since the priority of a packet is directly impacted by it. The COCR defines a range of different Classes of Service (CoS) which also refer to the priority of a packet (e.g. measured in terms of TD_{95}). Proper scheduling techniques ensure that packets belonging to a higher priority service are also transmitted earlier over the link. The scheduling thus addresses the design requirements, which state that the different services within a service category (i.e. for instance DG-C and DG-E within the ATS service category) can be prioritized. Furthermore the scheduling ensures that the different service categories can be prioritized among each other, i.e. for instance ATS over AOC. Finally the scheduling has a significant importance to ensure that OP services (i.e. ATS and AOC) are always prioritized over NON-OP services (i.e. AAC and APC). Since the queue size is in reality always limited, situations can occur where the buffers overflow, e.g. in situations where the link rate is lower than the arrival rate, the buffers fill up and finally overflow. In such a situation where the buffer is full but new packets arrive a decision has to be made on which packet needs to be discarded. There are three basic and intuitive possibilities:
  - Drop a random packet in the queue
  - Drop the packet at the first position in the queue
  - Drop the packet at the end of the queue (tail dropping)

In the context of OP services, the queue management policy may improve the QoS. Here applying a tail dropping policy is not necessarily a good approach, for instance in situations where a packet further in front in the queue is already outdated (e.g. due to a long waiting time in the queue) and the later arriving packet already contains the most recent information. In the case of applying a drop-tail policy the packet with the recent up-to-date information would be dropped whereas the previous packet with the outdated information is sent, since it is already in the queue. This is contra-productive to the goal of providing timely information. On top of this the interaction with higher layer transport protocols such as TCP is relevant. For instance dropping the first packet in the queue may trigger the TCP congestion avoidance algorithm already earlier (which is beneficial), but on the other hand may introduce unnecessary retransmissions

of later packets (which is undesirable). For this reason the selection of queuing policies is of particular interest for OP services when deploying a network.

Additionally, queuing policies try to address the issue of congestion control by applying so called active queue management (AQM). Here the queue length is continuously measured and, when exceeding a threshold, incoming packets are marked (to indicate an imminent congestion situation) according to a probability which is a function of the queue length or are directly dropped with this probability. The original purpose of this AQM was to support the behaviour of TCP and avoiding catastrophic congestion.

- Link selection strategies / routing decisions. Within the future aeronautical communication network, it is expected that many aircraft will have more than one data link technology. Besides legacy links such as the VHF based VDL-2, new link technologies, named Future Radio Systems (FRS) in COCR terminology, will arise. Examples for FRS are Aeromacs, LDACS, or future satellite communication links. For exchanging data a decision has then to be made which of the available links shall be used for transmission. The decision which link is favorable for the data exchange can depend on several criteria, such as cost of link usage, time before outage (e.g. due to leaving the coverage area or a handover), provided QoS and regulatory policies. The link selection strategy must on one hand collect information about the status of the different links and on the other hand try to find the best possible selection which is compliant with the requirements while at the same time minimizing the cost.

2.3.1 Separation of Operational and Non-Operational Domains

From today's perspective, one of the major design constraints is the demand for strict separation of operational (OP) and non-operational (NON-OP) services. This separation can be achieved on different layers:

- Separation at physical layer: Most rigorous form of separation. Here the OP and NON-OP services use different radio frequencies (RF) for transmission and remain entirely separated throughout the protocol stack up to the application layer.
- Separation at link layer: OP and NON-OP services use the same physical RF. Separation is achieved here by means of Link Layer segments, e.g. restricting that within a GSE Layer 2 cell only fragments of OP or only of NON-OP packets must be encapsulated.
- Separation at network layer: OP and NON-OP services may use the same physical RF frequency and also share Layer 2 cells. The separation is achieved here by different IP datagrams which are not shared among operational and non-operational services.

Fig. 6 illustrates the separation between OP and NON-OP service domains as expected for the near future.

As can be seen here, the domains are entirely separated down to the physical layer. The ATC and AOC services are connected to one mobile router, whereas the AAC and APC are connected to a different one. The strict separation of operational and non-operational services has far ranging consequences on the QoS architecture, especially with respect to Connection Admission Control (CAC), Congestion Control (CC) and traffic shaping as was explained earlier. The architecture shown in Fig. 6 was considered as the expected near term situation within the NEWSKY Project (NEWSKY, 2009).

In contrast to this, the more visionary approach which is also investigated in SANDRA is to have a full integration of different service domains into one network and to provide the

needed safety and security among OP and NON-OP services by means of networking techniques.
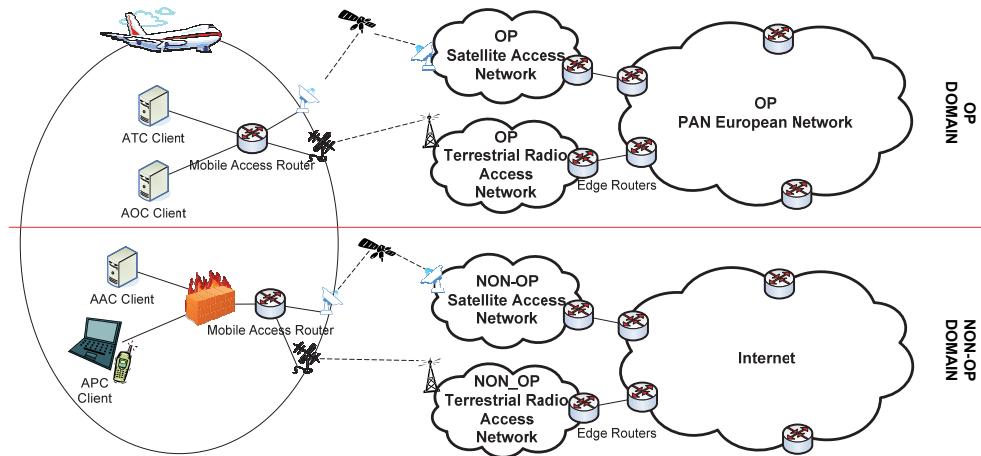
Fig. 6. Service domain separation for near future.

Fig. 7 illustrates an intermediate integration, where on the airborne side OP and NON-OP services are integrated. The division into OP links and networks and NON-OP links and networks still exists here.
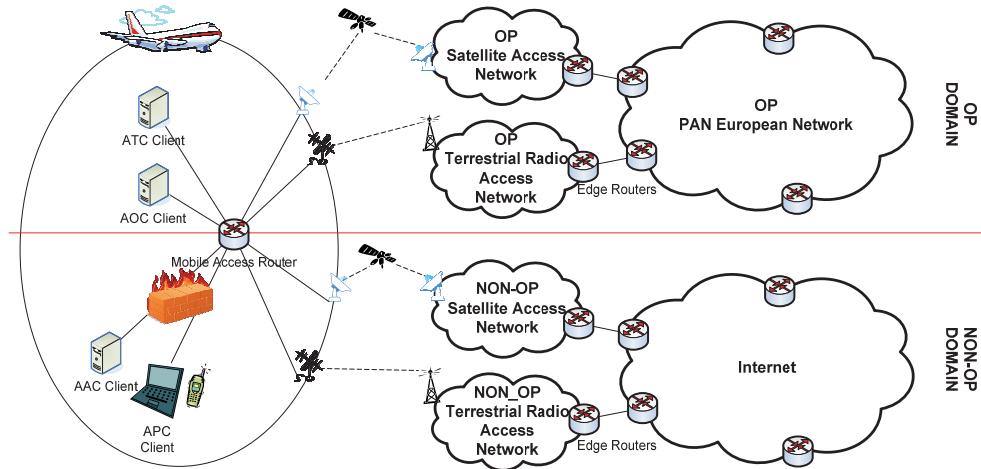
Fig. 7. Service integration in Mobile Access Router, separation at network domain level.

Here besides saving the additional equipment on board of the aircraft (the mobile router for the NON-OP domain) in principle the mobile access router has the freedom to route data over the same links or restrict due to policies the usage of some links, e.g. restricting the use of OP certified links for transporting NON-OP data. As long as the OP access networks on

ground are not interconnected with the NON-OP domain, sharing links between OP and NON-OP services is of course not very meaningful.

The relevance of the integration gets even more clear when looking into a fully-integrated scenario as shown in Fig. 8.
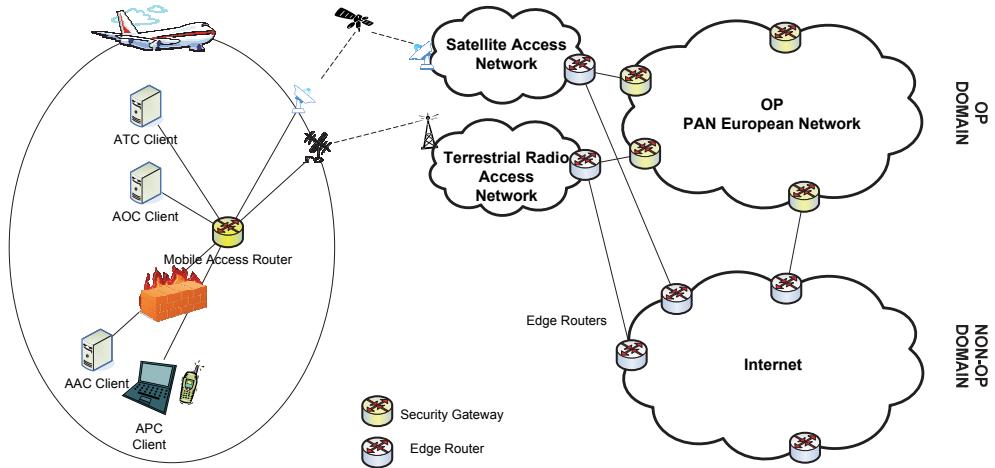

Fig. 8. Full integration of services, links and networks.

In this case the available links (SatCom and terrestrial radio in Fig. 8) may transport OP and NON-OP applications. The edge routers of the access networks then route the data to the other core networks, i.e. the OP PAN European Network domain or the public Internet. The edge routers of the OP PAN European Network domain additional have to provide security functionalities to avoid intrusion and corruption of incoming data. In principle a direct connection of the PAN European Network and the public Internet is conceivable, but not necessarily existing. It is clear that such an architecture creates a strong demand for strong and safe security mechanisms to protect the OP network, otherwise such an architecture will remain unacceptable due to safety concerns; as of now it is disallowed by regulation.

2.3.2 Underlying QoS approach

For provision of QoS different approaches are known from the literature. The suitability of the most well known ones, Integrated Services (IntServ) and Differentiated Services (DiffServ) for application in the aeronautical scenario is briefly reviewed in the following.

2.3.3 IntServ QoS approach

The IntServ architecture (Wroclawsky & Braden, 1997), (Zhang *et al.*, 1997) was developed for supporting specific QoS for end-to-end sessions across networks. In this approach, single flows (representing a stream of packets) are identified and treated individually. Every packet is checked for the resources it is entitled to receive. For this purpose the state of all flows in the network has to be periodically signalled among the routers in the end-to-end path of each flow. The Resource ReSerVation Protocol (RSVP) (Zhang *et al.*, 1997) was designed for this purpose. IntServ also has connection admission control mechanisms as an

integral part of its functionality which admits new traffic to the network only if sufficient resources are available. By doing all this IntServ can guarantee hard upper bounds for packet delays and packet loss caused by buffer overflow. Moreover IntServ can rely with RSVP on an existing and well deployed signalling protocol. The per-flow treatment also allows Multi-Level-Priority-Preemption (MLPP) which can be beneficial to differentiate ATM messages according to their priority and urgency.

While these IntServ features match very well with the QoS requirements in the ATM environment, the application of IntServ would have several major drawbacks. As is the case for all IntServ architectures, the main drawback is the scalability of the system and the signalling overhead. The traffic profile of ATM message exchange as predicted in the COCR consists of mainly small messages in the order few bytes, reaching at maximum several kilobytes in single cases. In the downlink for instance (i.e. aircraft to ground in ATM terminology) the maximum message size is 2763 bytes for the FLIPINT service. Estimations on the traffic profile have shown that the maximum message arrival rate hereby is slightly below 1 msg/s per aircraft at maximum, having an average of less than 0.1 msg/s per aircraft. This means in practice that either for every message a dedicated IntServ flow would have to be initiated and signalled, or an IntServ flow needs to be setup and kept alive for a longer time without being used most of the time, and accepting the overhead caused by the periodic keepalive messages necessary for this. Besides the volume overhead of the IntServ signalling also the time required for session initiation is an important overhead, considering that some messages have latency requirements as low as 0.74 s (Class of Service DG-B) and 1.4 s (Class of Service DG-C). For GEO satellite links already the session initiation would consume a considerable fraction of the maximum latency. Finally the heterogeneous and highly mobile environment, consisting of different link technologies and the belonging different access networks and the need for intra- and inter-technology handovers causes path changes. A change in the end-to-end path would then result also in additional IntServ session re-establishment overheads.

2.3.4 Differentiated Services (DiffServ)

DiffServ (Nichols *et al.*, 1998), (Blake *et al.*, 1998) is the second well known QoS architecture specified by the IETF. In contrast to IntServ no individual flows can be distinguished but only different aggregated classes of traffic. Instead of a guaranteed forwarding behaviour for every flow, DiffServ defines the per-hop forwarding behaviour for the aggregate classes. For identification of the aggregate, the Traffic-Class field in the IPv6 headers are used. Since in DiffServ only traffic aggregates are treated instead of single flows, no hard guarantees for the availability of resources and the end-to-end QoS performance can be given. An overdimensioning of resources is thus necessary here in order to meet the QoS requirements. The overdimensioning affects for instance the buffer sizes in the schedulers to avoid packet drops due to buffer overflow but also the available datarates on the links. While in theory the definition of one DiffServ aggregate per COCR Class of Service (CoS) would be possible (resulting in 12 aggregates), in practice a smaller number of DiffServ aggregates improves the scalability and reduces the complexity. In this case the application CoS need to be mapped by a classifier into the suitable DiffServ aggregates. Since all COCR CoS have different demands for maximum latency, an aggregation into fewer DiffServ aggregates implies also an increase of the required bandwidth, since the latency of the most demanding service in a DiffServ aggregate has to be met since DiffServ is not distinguishing within an aggregate. In other words services which could tolerate a longer latency need to

be transmitted in fewer time (i.e. the time of the most demanding service) what results in a higher demand in terms of data rate. For a DiffServ QoS approach also appropriate estimation and dimensioning of the network capacities is essential and requires a good model for the prediction of the amount of traffic to be transported including an additional buffer for unexpected traffic bursts. Such an (over)dimensioning on the other hand can also mean a waste of resources if capacity is strictly allocated per aggregate class and cannot be shared among different aggregates and considering the highly bursty traffic profile.

On the other hand a DiffServ architecture has significant advantages over an IntServ approach which outweigh the aforementioned drawbacks. Most important of all the issues with scalability do not exist here since only aggregates have to be treated instead of single flows. DiffServ is such much more suitable for the highly populated global ATM network under consideration with respect to this. Moreover a change of the end-to-end path, as can happen due to intra- and inter-technology handovers in this highly mobile scenario is not an issue here since no re-establishment of the RSVP tunnels is required anymore. Also the signalling overhead of IntServ for session initiation and keepalive can be saved while saving also the time for flow establishment which is beneficial for the overall delay profile.

2.4 Flow Identification

As was shown in other work (NEWSKY, 2009), routing decisions should be taken per flow, not per packet, e.g. due to problem of different latencies when messages are sent over different links, passing of packets, impact on TCP retransmission mechanism and reordering as well as load oscillations.

To identify the flow that a Layer 3 packet belongs to, the flow session identifier shall check the 5-tupel consisting of the IP source and destination address, source and destination transport layer ports and transport protocol.

In contrast to IPv4, which only allowed the identification of a traffic aggregate by the DSCP field or a particular flow, indicated by the 5-tupel, IPv6 additionally allows marking of single or aggregate flows via the flow label header field. Since also safety critical messages need to be exchanged in the aeronautical scenario, also security mechanisms such as IPSec may be applied. While encrpytion (IPSec Encapsulated Security Payload) may not be applied in all cases, means for authentication (IPSec Authentication Header) may be present. Considering the possibility to use IPSec also in tunnel mode, the flow identification can be done based on either inner or outer header (w.r.t. the tunnel) and before or after IPSec processing. Fig. 9 shows IPSec ESP tunnel mode for IPv6 datagrams.
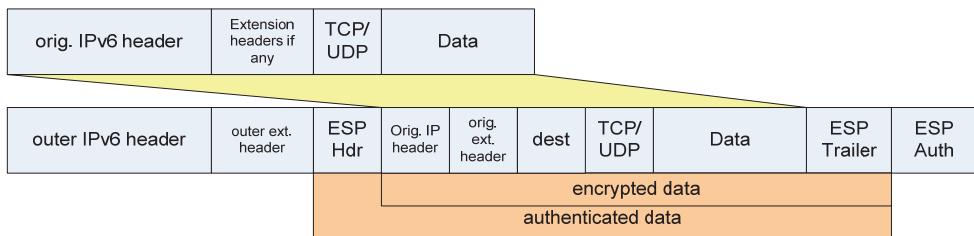

Fig. 9. IPv6 IPSec in ESP-tunnel mode.

In IPSec tunnel mode the inner header fields are not accessible in ESP mode since they are encrypted. Identification of the 5-tupel is not possible in these cases since also the UDP and

TCP headers, which are part of the 5-tupel, are located in the encrypted part. Though encryption is currently not envisaged for operational messages it is beneficial to do the flow identification before the IPSec processing since here identification of the 5-tupel can be done in any case.

In the case of dedicated Security Gateways (SG), the flow label assignment in the inner header must be done there, since after processing by the SG inner header fields must not be changed anymore. In case the SG is not implementing flow classification abilities, the flow label identifier in the router can only do a classification in case the inner header fields are visible (i.e. not encrypted) and only assign a flow label to the outer header.

In case the inner header fields are not visible no flow identification based on the original 5-tupel is possible.

For IPSec tunnel endpoints in the end systems (ES), it is the ES responsibility to set the correct values of the traffic class and flow label. As in the case of dedicated SGs, the subsequent routers can only do a classification in case the inner header fields are visible and flow label assignment can only be applied to the outer header fields.

The flow identifier also has to assign packets coming from the non-operational domain (AAC/APC) accordingly for a non-operational flow so the routing decision functionality can treat these packets seperately. The differentiation between operational and non-operational domain can be accomplished either IP address based or based on the physical interface:

IP address

In this case the OP and NON-OP traffic is distinguished only by the 5 tupel in the packet headers. This is however imposing a risking for spoofing attacks where these header fields are malicously modified by an attacker.

Physical interface

In this case the IR has different physical connector interfaces to the OP and NON-OP domain. Due to the physical separation, it is ensured that NON-OP data can in no case interfere with OP data, since a NON-OP packet is always unambiguously identified and treated.

For assigning the correct aggregate class, the flow identifier additional needs management information in form of DiffServ tables to map packets correctly to code points and flows IDs. These tables are specified in the management plane and allow configuration of the mapping

3. Conclusions on QoS architecture

In summary the following observations for the QoS architecture in an ATM can be made from the aspects briefly presented before:

A flow-oriented architecture such as IntServ would have the feature of guaranteeing a certain end-to-end behaviour, but is not suitable w.r.t. the bursty traffic profile, having only spurious transmission of single messages which have also only small size. The signalling overhead is considerable w.r.t. the small message payloads and also the additional time demand for a session initiation is considerable w.r.t. the latency requirements. A flow-oriented QoS architecture such as IntServ is thus no preferable solution for application in an ATM.

The alternative QoS architecture matching better with the given scenario is thus DiffServ. For deployment of a DiffServ QoS architecture several design parameters have to be kept in

mind, in particular the correct dimensioning of the resource trunks, mapping of application CoS into aggregate classes and priority scheduling. The main benefits here are the scalability also for a large and global ATM network. Also a change in the network point of attachment, e.g. due to a handover are not an issue here. The data volume and signalling delay overheads of IntServ can be saved here as well. For an integration of operational with non-operational services in the same network, however further specification of the mechanisms ensuring a safe separation of these two domains is required as well as deployment of mechanisms for CC, CAC and flow control of the NON-OP services.

Independently of the selected QoS model, the aeronautical QoS framework requires a solid and mature signalling framework, which can be easily derived from the experience acquired in ETSI BSM and IEEE 802.21 standardisation bodies. In particular, the extension of the SI-SAP primitives to match the aeronautical service requirements and the IR/IMR interaction are expected to be promising to help develop a fully QoS-oriented aeronautical architecture. On the other hand, the joint use of the aforementioned ones and the MIH framework should also guarantee an important support to efficiently manage the available transmission links and perform their selection accordingly.

4. Acknowledgments

5. References

Ali, M.; Xu, K.; Pillai, K. & Hu Y. F (2011). Common RRM in Satellite-Terrestrial based Aeronautical Communication Networks. In: *Third International ICST Conference on Personal Satellite Services (PSATS 2011)*, Malaga, Spain, February 2011

Blake, S.; Black, D.; Carlson, M.;. Davis, E.; Wang, Z. & Weiss, W. (1998). RFC 2475: An Architecture for Differentiated Services, *IETF*, 1998

EUROCONTROL (2007), FAA, Communications Operating Concept and Requirements for the Future Radio System (COCR), EUROCONTROL/FAA, 2007

Hitoshi, S. (1998) Teletraffic Technologies in ATM Networks, Artech House.

ICAO (2009), International Civil Aviation Organization: Manual for the ATN using IPS Standards and Protocols (Doc 9896). February 2009, 1st edition, Unedited Advance version

ITU-T Recommendation G.1010 (2001) End-user multimedia QoS categories, Nov. 2001.

Marchese, M. (2007), QoS over Heterogeneous Networks, *John Wiley & Sons Ltd*, 2007.

NEWSKY (Networking the Sky for Aeronautical Communications) (2009) – www.newsky-fp6.eu

Nichols, K.; Blake, S.; Baker, F. & Black, D. (1998) RFC 2474: Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers, *IETF*, 1998.

Plass, S.; Kissling, C.; De Cola, T. & Van Wambeke, N. (2011): The SANDRA communications concept - Future aeronautical communications by seamless

networking. In: Proc. Third International ICST Conference on Personal Satellite Services (PSATS 2011), 17.-18. Feb. 2011, Malaga, Spanien.

SANDRA (Seamless Aeronautical Networking through integration of Data links Radios and Antennas) (2011) – www.sandra.aero

Tannenbaum, A. (2002) Computer Networks, Prentice Hall Professional Technical Reference, 2002.

Wroclawsky, J. & Braden, R. (Ed.) (1997). RFC 2210: The Use of the Resource Reservation Protocol with Integrated Services, *IETF*, 1997.

Zhang, L.; Berson, S.; Herzog, S. & Jamin, S. Braden, R. (Ed.) (1997). RFC 2205: Resource ReSerVation Protocol, (RSVP) --- Version 1 Functional Specification, *IETF*, 1997

Interoperability Among Heterogeneous Networks for Future Aeronautical Communications

Kai Xu, Prashant Pillai, Yim Fun Hu and Muhammad Ali
School of Engineering, Design and Technology, University of Bradford
United Kingdom

1. Introduction

Air traffic in Europe is expected to double by 2025 according to the last forecast of EUROCONTROL, with an average growth of 2.7%-3.7% per year. On a worldwide basis, the number of passengers is expected to grow by 4.5% per year over the same timeframe. Future passenger and freight fleets will bring higher efficiency and improved environmental performance, and will allow people all around the world to benefit from the essential connections that only air transport can deliver. (European Organisation for the Safety of Air Navigation [EUROCONTROL], 2006)

Till recently Aeronautical Telecommunication Network (ATN) communications predominantly across continental regions used narrowband Very High Frequency (VHF) voice systems along with digital data links like the VHF Digital Link (VDL) Mode 2 (ICAO, 2001) and Aircraft Communications Addressing and Reporting System (ACARS) (ARINC, 2011). Such narrowband air-ground data link technologies have several limitations like the susceptibility to access collisions, lack of data prioritization mechanism at the sub-network level and also the vulnerability of jamming in narrowband communication channels. Hence, other High Frequency (HF) radio link technology like the ARINC GLOBALink (Homans, 2002) and also satellite based system like the INMARSAT SwiftBroadband systems have been used to provide voice and data communications between the aircraft and the ground control centres (EUROCONTROL, 2006). A detailed study on the various communications systems suitable for Air Traffic Management (ATM) systems was carried out by the Federal Aviation Administration (FAA) and EUROCONTROL with the support of NASA (NASA, 2005). It highlighted the advantages and disadvantages of the various communications systems and showed that it is important for the future ATN communication systems to provide fast, high-speed and high-data rate reliable communications not only between the aircraft and ground infrastructure but also between airplanes directly. Hence it is envisaged that in order to provide the most efficient form of reliable ATM communications, different heterogeneous networks need to be adopted. It is also envisaged that future aeronautical communication systems would be able to simultaneously operate the different radio communication technologies. The different technologies will provide various advantages and provide the flexibility to the system to provide very high reliability. Hence different data i.e. Air Traffic Service (ATS), Airline Operation Centre (AOC), Airline Administrative

communication (AAC) and Aeronautical Passenger Communications (APC) can be sent over different radio technologies depending on the Quality of Service (QoS) and security requirements of the data.

2. Wireless digital data links for future aeronautical communications

The growing need to efficiently manage airspace for increased traffic loads will require new and more capable communications systems. The joint EUROCONTROL/FAA technology assessment study, known as the Action Plan 17 - Future Communications Study (FCS), made specific recommendations for data communications technologies in VHF, C, and L bands that were endorsed by the FAA, EUROCONTROL, and International Civil Aviation Organization (ICAO). (EUROCONTROL/FAA, 2007)

2.1 VHF

In the aeronautical communications, Air Traffic Control (ATC) is a very important service provided by ground-based controllers who direct aircraft on the ground and in the air. VHF radio is used by the Ground Control for monitoring and controlling any aircraft, vehicle or person walking or working in the specific areas, such as taxiways, runways and departure gate, to have clearance. This requires most aircrafts to have VHF radios. VHF is the radio frequency ranges from 30MHz to 300MHz. Common uses for VHF are radio/television broadcast, ATC communications and air navigation systems.

The VHF Digital Link (VDL) communications system is one of aircraft-to-ground subnetworks that may be used to support data communications across the ATN between aircraft-based application processes and their ground-based peer processes. The digital communications protocols are employed on the devices using VHF radios to support data communication functions. The International Telecommunication Union (ITU) assigns the band 118MHz to 137MHz to be used by VDL for the aeronautical services. A range of International Civil Aviation Organization (ICAO) standards are defined by Aeronautical Mobile Communications Panel (AMCP):

- ICAO VDL Mode 1: is defined for validation purposes. The VHF analog radio is used before VHF Digital Radio implementation was completed.
- ICAO VDL Mode 2: is the main version of VDL. This is the only mode being implemented operationally to support Controller Pilot Data Link Communications CPDLC (i.e. an ATC service). The VDL2 system is based on the OSI reference model, therefore the lower three layers: the physical layer, the data link layer and the subnetwork layer, are specified to provide code transparent communications between the ATN systems. The physical layer provides services to activate, maintain and de-activate connections for bit transmission in the data link layers. The link layer is responsible for transferring information from one network entity to another. The subnetwork layer is responsible for controlling the data packet flow with respect to duplicate, lost or invalid packets. A Subnetwork Dependent Convergence Function (SNDCF) is required to provide a transparent connectionless mode service to the ATN internetworking protocol, as the VDL2 is a connection-mode subnetwork access protocol. (ICAO, 2001)
- ICAO VDL Mode 3: defines a protocol providing aircraft with both data and digitized voice communications that was defined by the US FAA.

- ICAO VDL Mode 4: specifies a protocol enabling aircraft to exchange data with ground stations and other aircraft.

From the aeronautical mobile communications evolution overview carried out by, the VHF analog voice communications will be maintained and operated throughout the time frame (2005 - 2030) both in the European and in the U.S. EUROCONTROL is progressing toward the full implementation the carriage and operation of 8.33 kHz capable equipment below FL195 in 2013. In addition to legacy ACARS services, Europe will use VDL Mode 2 in the VHF band for the ATC and AOC data link services. Air traffic management services that operate on VDL Mode 2 will be maintained throughout the time span, enhanced as needed to support safety related services in the U.S. (EUROCONTROL/FAA, 2007).

2.2 DVB-S2

Digital video broadcasting-satellite-second generation also known as DVB-S2 is a successor for the well-known digital video broadcast standard DVB-S. The DVB-S2 is based on Digital Satellite News Gathering (DSNG) and DVB-S standards used by mobile stations for transmitting multimedia data to their home television stations from remote locations all over the world. The DVB-S2 added two new and significant features which are: A modern Low-Density-Parity-Check (LDPC) powerful coding scheme and secondly a Variable Coding & Modulation (VCM) and adaptive coding & modulation (ACM) parameters which enable dynamically changing transmission parameters to optimize bandwidth utilization. Additional features which DVB-S2 provides are improved modulation schemes up to 32APSK (ETSI, 2009b), generic transport mechanism for IP packet data also including MPEG-4 video and audio streams support as backward compatibility with DVB-S MPEG-2 TS based transmission and additional code rates. The DVB-S2 standard boasts that DVB-S2 has almost 30% performance gain over DVB-S using the same satellite transponder bandwidth and emitted signal power. HDTV service can now be delivered using the same capacity which supported earlier DVB-S based SDTV service using the improved techniques of video compression in DVB-S2. Several technological advancements have been achieved in the last decade for satellite broadcasting, which has given rise to the need of DVB-S2. (ETSI, 2009b)

The DVB-S2 standard attains high data transmission capacity as compared to first generation system (DVB-S) as DVB-S2 has less complex design for hardware and software level implementation of receiver, provides an enhanced flexibility. DVB-S2 takes advantage of the recent advancement in modulation and channel coding techniques to maintain equilibrium between performance and complexity. Implementation of such novel techniques to achieve target goals, enable DVB-S2 to raise 30% capacity as compared to DVB-S by utilizing the same transmission conditions. The DVB-S2 can provide Constant Coding and Modulation (CCM) and Adaptive Coding Modulation (ACM). (ETSI, 2009b)

The DVB-S2 enables high flexibility to deal with characteristics of all existing satellite transponders with a large variety of spectrum efficiencies and associated C/N requirements. Supporting format for DVB-S2 is not only limited to MPEG-2 multimedia data source coding but it is efficiently designed to deal with a variety of video, audio and data formats including the formats which DVB projects is currently developing for the future application. The DVB-S2 system is optimized for the broadband application scenarios such as: Interactive services. The internet access can be a good example of interactive service. DVB-S2 is envisioned to provide the interactive services especially to the mobile devices roaming

in the remote places. As DVB-S and DVB-S2 both provide the transmission in forward direction only there is no detailed mechanism specified in the standard documents of DVB-S and DVB-S2 about the return direction. Therefore the interactive services can be established using return link as terrestrial or other satellite links. Generic stream format or transport stream format are utilized for data services transportation (ETSI, 2009b).

2.3 BGAN

In December 1999, INMARSAT's Board of Directors approved the development of fourth generation of satellites for the Broadband Global Aeronautical Network (BGAN) project. The INMARSAT-4 satellites comprise of two GEO satellites, which were initially located over the Indian Ocean Region (64ºE) and Atlantic Ocean Region (53ºW) respectively, to provide communication service to the Mobile Terminals (MT). A third satellite (98ºW) was launched in 2008. Each INMARSAT-4 satellite weighs 3 tonnes and supports approximately 200 spot beams. It provides transparent amplification for the BGAN communications (user plane and control plane). Transmission between the RNC and satellite is via the C band, whereas transmission between the satellite and MTs is via the L band. (INMARSAT, 2003)

The INMARSAT BGAN is intended to form part of the satellite component of the Third Generation (3G) IMT-2000/Universal Mobile Telecommunications System (UMTS). Among the design objectives is the interoperability with an industry standard 3G Core Network and the re-use of the UMTS Non Access Stratum (NAS) layers. The BGAN system adopted the standard UMTS architecture in the core network but uses an INMARSAT proprietary air interface –the INMARSAT Air Interface 2 (IAI-2), which provides a complete Access Stratum Protocol Stack and Physical Layer optimised for the geo-stationary satellite environment. The air interface is based on TDM and TDMA/FDM schemes in forward direction and return direction respectively (INMARSAT, 2003).

BGAN is the first system to provide guaranteed data rates on demand. It is also the first mobile communication system to provide both voice and broadband mobile communication services on a global area, where two GEO satellites of BGAN system will cover 85 percent of the earth's surface. The system network architecture has the capability to provide UMTS compatible services and both circuit switched and packet switched services. The BGAN MTs are the lightweight, portable satellite terminals, which are easy to carry, simple to setup for use, can deliver data rates of up to half a megabit. In order to achieve high transmission efficiency and flexibility, it is possible to adapt the bandwidth, coding rate according to the MT's class and channel conditions. In the BGAN baseline system, 3 classes of MTs are supported with different maximum transmission rates of 432 Kbps, 432 Kbps, and 216 Kbps respectively when receiving data and maximum transmission rates of 432 Kbps, 144 Kbps and 72 Kbps respectively when transmitting. All MTs have the capability of allocating bandwidth dynamically in each direction. At present, the number of MT classes is being extended to support additional services in the BGAN-X project. (Jilg, 2002; ESA, 2003; INMARSAT, 2003)

BGAN constellation aims to support existing aeronautical safety services and INMARSAT has made a Public Service Agreement (PSA) commitment to ICAO. In March 2006, Thrane & Thrane, the world leading INMARSAT satellite communications company, entered a BGAN Extension contract with INMARSAT. The contract covers an upgrade of the RAN Satellite Access Stations to also handle the planned broadband services for aeronautical (SwiftBroadband) usage in providing voice and data services.

2.4 AeroMACS

The Aeronautical Mobile Airport Communications System (AeroMACS) is a new airport communication system. This system is jointly identified and recommended based on IEEE 802.16e system by the EUROCONTROL and American Federal Aviation Administration (FAA) to provide the solution for dedicated aeronautical communication services on the airport surface. The EUROCONTROL and American Federal Aviation Administration (FAA) have established the first Memorandum of Cooperation (MoC) in 1986 (EUROCONTROL, 2001). To accommodate future growth in surface applications, the new airport radio local area network should operate in the additional spectrum in C-band (5091 – 5150 MHz), which is co-allocated in the World Radio Communications Conference 2007 (WRC07). It is identified that the AeroMACS radio technology is based on the IEEE 802.16e standard (IEEE, 2009a), which is best suitable for short-range and mobile broadband data communications. The AeroMACS technology is well matched to the airport surface environment in terms of capability and performance. The AeroMACS standards development activities, such as to propose an aviation specific standard and to evaluate/validate the performance of the standard etc., have been planned under the EUROCONTROL/FAA AP 17. (EUROCONTROL/FAA, 2007)

As the development of AeroMACS is based on the IEEE 802.16e standard, it is believed by EUROCONTROL that no (or minimum) modifications or additions will be needed to the existing IEEE organisations' profiles according to the aviation special operational requests. But EUROCONTROL is still looking into the need for establishing new system profiles which shall be optimised for aviation needs. The existing IEEE profiles are split up into two parts: System Profiles - specifying subsets of mandatory and optional physical and MAC layer features from the standard and the Certification Profiles - specifying the channel bandwidth, operating frequencies and duplexing method. Because the AeroMACS radio will operate in dedicated aeronautical spectrum (5091 – 5150 MHz), it is need to establish a new Certification Profile. (EUROCONTROL, 2009)

IEEE 802.16e standard is the most popular commercial implementation of the 802.16 family of standards, which is written by a working group established by IEEE Standards Board. While WirelessMAN is the official name of the 802.16 family, it is commercially known as Worldwide Interoperability for Microwave Access (WiMAX). The name is coined by the WiMAX Forum, which is an industry alliance to promote and certify compatibility and interoperability of broadband wireless products based on the IEEE 802.16 standards. IEEE 802.16e (abbreviation of 802.16e-2005) concluded in 2005, specifies mechanisms to support mobility. It is therefore also known as Mobile WiMAX. Because of its advanced transmission capacities of between 34 Mbits/s and 1Gbit/s, WiMAX networks are able to provide mobile broadband or at home broadband connectivity across whole cities or countries.

3. Interoperability among heterogeneous networks

As previously discussed, it is envisaged that future ATN system will comprise of heterogeneous systems. Hence it would be important to monitor and control these various systems efficiently and also provide means to handover active sessions between these different radio technologies. Currently, the different radio access networks work independently and have limited or in most cases no interaction amongst themselves. However, to support efficient control and seamless vertical handovers, future networks need to interact and communicate together. Recently, some standardisation work has been

carried out in order to define mechanism and protocols to support interworking between different networks.

The following subsections provide an overview of two existing standards, namely the ETSI Broadband Satellite Multimedia (BSM) SI-SAP (Satellite Independent – Service Access Point) and the IEEE802.21 Media Independent Handover framework that provide means to support interoperability among heterogeneous networks. Both standards consider the separation of technology-independent upper layers from the technology-dependent lower layers to enable interoperability among heterogeneous networks. The BSM SI-SAP concentrates on the separation of upper layer functions from lower layer functions for satellite systems whereas the MIH framework defines media-independent functions for handover between heterogeneous mobile and wireless networks.

3.1 Broadband satellite multimedia

The BSM architecture standardised by ETSI, defines a Satellite-Independent Service Access Point (SI-SAP) interface (ETSI, 2005), which is presented as a radio abstraction layer in the BSM protocol architecture to separate the satellite-independent functions in the upper layers (layer 3 and above) from the satellite-dependant functions in the lower layers (layer 2 and below), thus providing a mechanism to carry IP based protocols over these satellite-dependent lower layers to provide efficient IP-based broadband services over heterogeneous satellite systems. The BSM reference architecture consists of three major groupings of BSM elements, as shown in Fig.1. These are the BSM System (BSMS), BSM Network and the BSM subnetwork. They correspond to a BSM Network together with the NMC and NCC plus any additional elements that are required to provide network services. The BSM Network corresponds to a BSM subnetwork together with BSM interworking and adaptation functions. Finally, the BSM subnetwork consists of all the BSM network elements below the Satellite Independent Service Access Point (SI-SAP). The SI-SAP is the common interface between any BSM family of satellite dependent lower layers and the satellite independent upper layers.
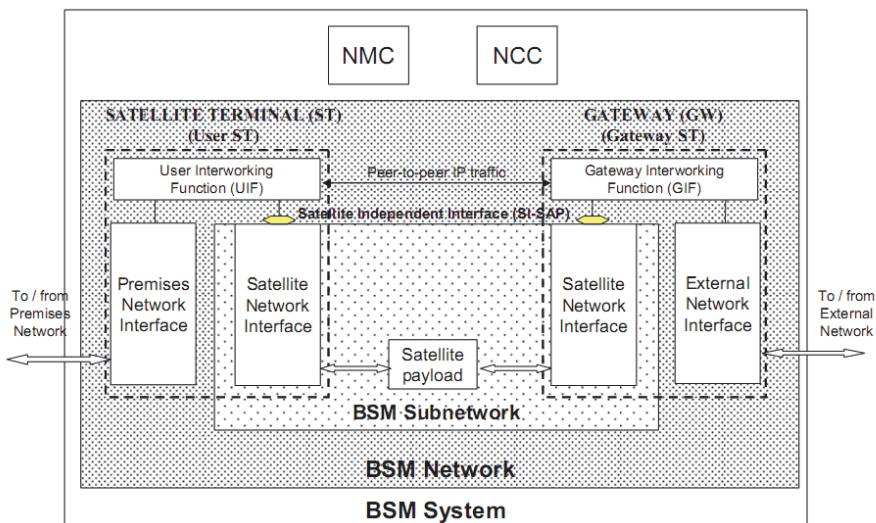

Fig. 1. BSM reference model (ETSI, 2007).

The BSM protocol architecture supports families of air interface protocols, where each family defines a complete stack of air interface protocols for the physical layer and the data link layer. Each air interface family is expected to use a combination of a Satellite Link Control (SLC), Satellite Medium Access Control (SMAC) and Satellite Physical (SPHY) layers that are jointly optimised for a specific range of satellite architectures and/or for a specific range of traffic type (ETSI, 2007).
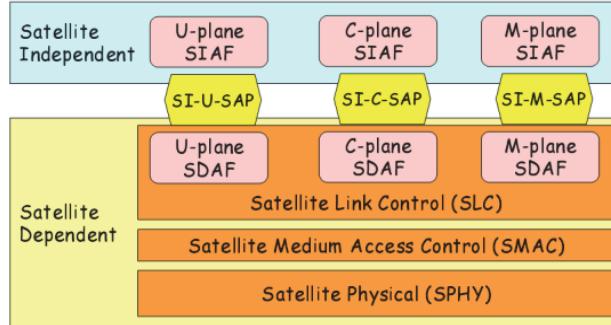

Fig. 2. SI-SAP expanded reference model (ETSI, 2007).

The SI-SAP reference model is shown in Fig. 2. From the diagram, it can be seen that the SI-SAP interface provides a standard interface to upper layers of the protocol stack irrespective of the satellite dependent lower layers. The Satellite Independent Adaptation Functions (SIAF) operate at the bottom of Layer 3 to adapt between the layer 3 protocols and the SI-SAP services, whereas the Satellite Dependent Adaptation Functions (SDAF) operate at the top of Layer 2 to adapt between the SI-SAP services and the native services of the given BSM family. SI-SAP and the associated adaptation function can be divided into the U-plane, C-plane and M-plane services. SI-U-SAP deals with services to transport IP packets between STs. The SI-C-SAP is concerned with control and signalling related to the user plane services, whereas the SI-M-SAP deals management services over BSM.

3.2 Media independent handover framework

The IEEE802.21 Media Independent Handover (MIH) framework (IEEE, 2009b) defines a unified interface between different link layer technologies for the support of seamless mobility between heterogeneous IEEE 802 networks and between IEEE 802 and other mobile wireless technologies. This unified interface is presented as an abstraction layer function, the Media Independent Handover Function (MIHF), for handover detection, initiation and decision via Layer 2 triggers.

Fig. 3 shows the IEEE802.21 MIHF reference model and SAPs. An MIHF in a network entity that communicates directly with an MIHF in a mobile node acts as a Point of Service (PoS) of that mobile node (MN). The MIHF receives media independent commands from higher layers and translate them to media specific commands for link layer and similarly receives events from different link layer technologies and maps them to corresponding media independent events. The MN exchanges MIH information with its MIH PoS using L3 transport if the PoS does not reside in the same network entity as its network Point of Attachment (PoA). A PoA is the network side of a layer 2 link that includes an MA as the other end point. Thus, the MIH framework supports both L2 and L3 transports for exchanging MIH information.
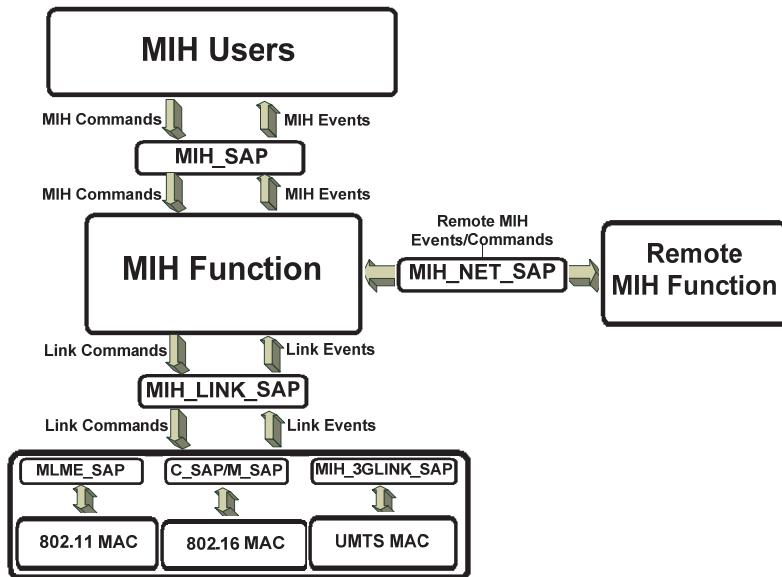
Fig. 3. General IEEE802.21 MIHF reference model and SAPs (IEEE, 2009b).

To facilitate media independent handover, the MIHF provides three services: Media Independent Event Service (MIES), Media Independent Command Services (MICS) and Media Independent Information Services (MIIS). The MIES reports events on dynamic changes in link characteristics, links status and link quality to upper layers through the MIHF. The MICS is used to gather information about the status of the connected links. Upon reception of event notification, MIH users make use of the MICS to pass link commands to the lower layers via the MIHF to manage and control the link layer behaviour for handover decision. MIIS provides the capability for obtaining the necessary information for handovers, including neighbouring networks, link layer information and service availability. This information will be used to assist network discovery and selection to enable more effective handover. Entities that use the services provided by the MIHF are called MIH users. MIH users access MIHF services through a variety of SAPs. Each SAP consists of a set of service primitives that specify the interactions between the service user and provider. Three SAPs are currently defined within the MIH framework: the MIH_SAP for the Upper Layer access to the Lower Layers via the MIH; the MIH_LINK_SAP to connect the MIH Function and the Lower Layers, and MIH_NET_SAP for service transport between the local and the remote MIHFs.

- MIH_SAP: A media independent interface provides the interface between the MIHF and the upper layers of the mobility management protocol stack. In order to receive MIHF generated events and link layer events that are forwarded by the MIHF, the upper layers need to subscribe with the MIHF as MIHF users. MIHF users can directly send commands to the local MIHF using the service primitives of the MIH_SAP.
- MIH_LINK_SAP: An abstract media dependent interface between the MIHF and media specific lower layer to allow MIHF to use services from the lower layers of the mobility management protocol stack. For each link layer technology, the MIH_LINK_SAP maps to the media specific SAPs.

- MIH_NET_SAP: An abstract media dependent interface of the MIHF, provides transport services over the data plane on the local node to support the exchange of MIH information and messages with remote MIHFs. Transport services provided by the MIH_NET_SAP can use either L2 or L3 signalling.

4. The SANDRA radio resource management (RRM) framework

4.1 SANDRA network architecture

The EU Project SANDRA (Seamless Aeronautical Networking through integration of Data-Links, Radios and Antennas) (SANDRA, 2011; Denos, 2010) aims to design, specify and develop an integrated aircraft communication system to improve efficiency and cost-effectiveness by ensuring a high degree of flexibility, scalability, modularity and reconfigurability. The SANDRA system is a "system of systems" addressing four levels of integration: Service Integration, Network Integration, Radio Integration and Antenna Integration. In addition to these four levels of integrations, an important aspect of SANDRA is the development of AeroMACS – a WiMAX adaptation for aeronautical communications - for integrated multi-domain airport connectivity. From the communications network point of view, SANDRA spans across three segments, namely, the Aircraft segment, the Transport segment and the Ground segment, as shown in Fig. 4.

The Aircraft segment consists of three main physical components: the Integrated Router (IR), the Integrated Modular Radio (IMR) and the Antennas. These three components form the SANDRA terminal. The IR is responsible for upper layer functionalities, such as routing, security, QoS and mobility. The IMR takes care of lower layer radio stacks and functions including radio resource allocation, QoS mapping and adaptation functions. Through Software Defined Radio (SDR) (ETSI, 2009a), the IMR supports dynamic reconfigurability of operations on a specific radio link at any time and provides the flexibility for accommodation of future communication waveforms and protocols by means of software change only.
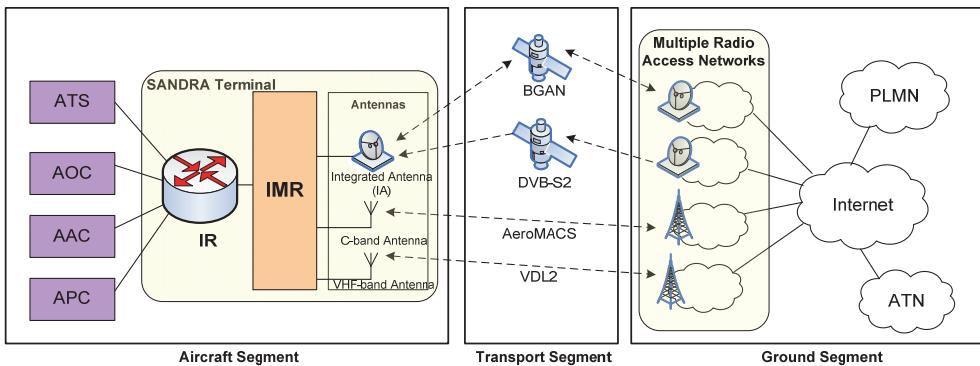

Fig. 4. SANDRA network architecture.

The physical separation between the IR and the IMR has the advantage of increased modularity and identifying distinct management roles and functions for higher layer and lower layer components with IP providing the convergence. The Antennas include a hybrid Ku/L band Integrated Antenna (IA), a VHF antenna and a C-band antenna. The IA is a

hybrid Ku/L band SatCom antenna to enable an asymmetric broadband link. The Ku-band system is used for receive mode only to reduce the number of antennas and the L-band will use the INMARSAT Swift Broadband link capable of both transmit and receive functions. The VHF antenna will be used to provide transmit and receive links for VDL2 mode ATC and AOC voice communications services. The C-band antenna will be used for AeroMACS to enable aircraft-airport and inter-domain airport connectivity. The various end-systems i.e. ATS, AOC, AAC and APC (SANDRA, 2011) are all connected to the IR.

In the transport segment, four radio transport technologies are considered, namely, VDL mode 2 (ICAO, 2001) in VHF band for legacy systems and applications, BGAN (INMARSAT, 2011) in L-band, DVB-S2 (ETSI, 2009b) in Ku-band for passenger communication services and AeroMACS (a WiMAX equivalent for aeronautics communications) in C-band providing airport link for short range airport ATC communications. These transport technologies provides connectivity between the aircraft segment and the ground segment supported by the IMR and their corresponding Radio Access Networks (RANs) on ground. The selection of the appropriate transport network is made within the SANDRA connection manager (CM) that spans across the IR and the IMR based upon applications, policies, regulatory constraints, service availability, quality of service and other factors.

The Ground segment consists of multiple operators; multiple Radio Access Networks (RANs) and their corresponding core networks, the ATN, the Internet and possibly the Public Land Mobile Network (PLMN, for passenger communications). The RANs can also be connected directly to the ATN and the PLMN on the ground. In order to provide mobility service and security services for aeronautical communications, functional components such as the mobility server, security and authentication server are required in the ground segment to provide corresponding mobility information services as well as security services. These components will be provided by the ATS/AOC/AAC and APC service providers of the ATN on ground.

4.2 Overview of RRM

The efficient utilisation of common radio resources is an important aspect in the wireless communication networking. The goal of Radio Resource Management (RRM) is to optimize bandwidth utilisation and Quality of Service, in the presence of traffic flows generated by services with different requirements. QoS can refer to the capability of a telecommunication system to provide better services to user traffic. Whenever resources or their modifications are requested, the goal of RRM is to optimise resource utilisation and at the same time satisfy the requested QoS requirements while trying to maintain a certain degree of fairness among all users. Cross-layer RRM techniques that consider the cooperation of two or more protocol layers have been used to optimise the system usage in run time to guarantee good performance of the overall system in order to fulfil the user performance requirements. (Giambene, 2007)

An efficient dynamic RRM scheme can optimise the system usage in run time to guarantee good performance of the overall system by maximising some performance indicators, such as the overall network throughput, the resource utilization and total network earning, and at the same time to minimise other indicators, such as the end-to-end (ETE) delay, the packet discard rate (PDR), the call dropping rate and the signal-to-noise ratio. RRM schemes can include a set of service control functions, which are categorised into network based functions and connection based functions. Network based functions include admission

control (AC), load control (LC), packet scheduler (PS) and resource manager (RM); whereas connection based functions include power control (PC) and handover control (HC):

- Admission control function is to grant/deny connections' access to the network based on whether there are sufficient free resources by taking into account predefined criteria (QoS requirements). The AC process occurs when new connection is set up, as well during handovers and service modification. It is especially important in congestion control and in QoS provisioning for wireless networks. May AC algorithms, like conventional, expert knowledge and intelligent predictive/learning algorithms, have been proposed to resolve radio resource allocation problem.

- Load control is responsible to optimise the capacity of the system and to prevent overloading situations. This function enables the estimation of the traffic load condition and load balancing mechanism to be carried out. It is particularly important in managing system overload situations (i.e. when the system load exceeds the threshold limit) in order that the system load can be maintained in a feasible manner.

- Packet scheduling handles all traffic generated by packet data users and decides the order of packet transmission, when a packet transmission is initiated and what bit rate and resource will be allocated. This provides multiplexing functions for different traffic mix, taking into account the QoS constraints such as bandwidth requirement and traffic priority. The nature of a scheduling framework can greatly impacts the QoS levels that can be provided in the system. Much research related to scheduling techniques has been done focused on the latest wireless technologies to support diverse QoS requirements for a wide range of applications, for example WiMAX network, Orthogonal Frequency-Division Multiple Access (OFDMA) wireless system and shared-medium wireless network etc (Fantacci et al., 2009; Kumar et al., 2009).

- Resource manager includes the estimation of the required number of transport and physical channels for the support of various traffic types and the actual radio channel allocation based on the traffic load condition. Many research works have been carried out to investigate suitable resource access strategies for efficient radio bandwidth assignment. Four primary strategies are identified: fixed assignment, demand assignment, free assignment and random access. Many hybrid resource allocation strategies have been proposed by combining the good advantages of the different strategies.

- Power control transmits the signal with the lowest possible power level, but at the same time, maintaining the required signal quality to maximise network capacity. Determining the transmitter power level is quite a challenging task due to the dynamic variation of the radio channel. To achieve this, power control functions have been well investigated and many efficient algorithms have been developed to accurately control the transmission power, such as centralised, distributed, synchronous, asynchronous, interactive and non-interactive methods.

- Handover control function provides the ability for the user devices to change their access methods to another network in order to maintain connectivity when their environment changes. Three main phases in handover: initiation, decision and execution, are processed to ensure call/session continuity. If a handover is performed between two heterogeneous networks, a handover at the network layer making use of Mobile IP is normally used to enable communication and signalling between the two networks through IP. An important aspect of handover control is to ensure that QoS

guaranteed by the serving network can be met by the target network (network to which the call/session is to be handed over). A mapping between the QoS parameters at the IP layer and those supported by the bearer services in the radio access networks has to be made.

4.3 Software defined radio

The components in a radio communication system are typically implemented in hardware. A Software Defined Radio (SDR) system implements the components instead by means of software computing device. The future of telecommunications is anticipated to provide great variety of services over a multitude of Radio Access Technologies (RATs). To achieve this aim, it is required that the future system is to support heterogeneity in wireless access technologies, comprising different services, device capabilities, and so on. The SDR technology can provide hardware reconfiguration ability and software programmability by introducing the open architecture to the programmable digital radio. The architecture is able to maintain an independency between the hardware function module and software function module. With the architecture, the SDR technology can reconstruct the application software in one common hardware platform in order to flexibly manager with the various wireless technologies with the purposes of increasing the frequency use efficiency and improving QoS. This technology is expected to form the new framework of the wireless telecommunications in the future.
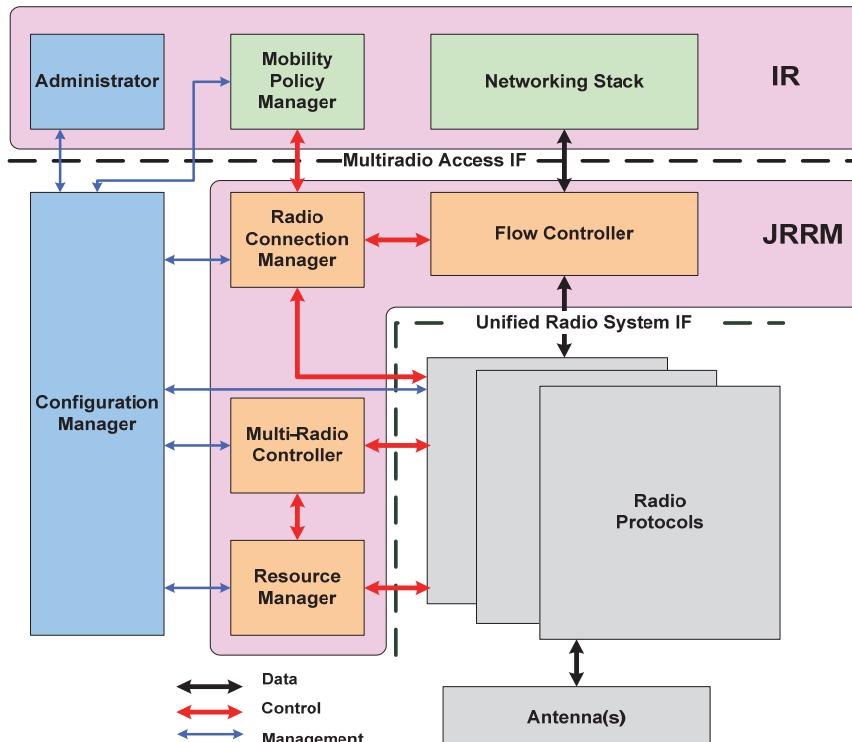

Fig. 5. The SDR-enabled RRM Functional Architecture for the SANDRA Terminal.

The Integrated Modular Radio (IMR) of the SANDRA terminal makes use of Software Defined Radio (SDR) technology to provide flexible management and reconfiguration of multiple radio access technologies supported by the SANDRA terminal in a common hardware platform. The functional architecture within the SANDRA terminal from purely the Radio Resource Management (RRM) perspective will be provided in the next section. Such architecture needs to be extended to encompass the SDR technology offered by the SANDRA terminal. Fig. 5 shows a possible SDR-enabled RRM functional architecture for the SANDRA terminal based on the SDR control framework defined in (ETSI, 2009b; ETSI, 2009c) consisting of the following:

- Functional components supported by the Joint Radio Resource Management
- A Uniform Radio System Interface (URSI)
- A Waveform Configuration Manager (WCM).

The Multiradio Access Interface is supported by a common SDR control framework - Joint Radio Resource Management (JRRM), which consists of common functionalities to all radio systems. In particular, the URSI enables every radio system to be integrated into the common hardware platform in a unified manner. The Radio Protocols and Engines component corresponds to the waveform applications for individual stacks and the common programmable hardware platform. With reference to the RRM functional architecture provided in the next section, it can be seen that the following interfaces and their associated primitives will be defined:

- Interface between the Network Stack and the flow controller – this is the SANDRA-U-SAP interfaces.
- Interface between the Flow controller and the Radio Protocols and Engines is equivalent to the R-U-SAP.
- The MAI provides a common interface between the IR and the IMR. This is based on the BSM SI-SAP primitives.
- The URSI provides the common API for each radio stacks and will carry the MIH-LINK-SAP primitives.

4.4 Collaborative RRM functional architecture
4.4.1 CRRM architecture overview

This section presents a functional architecture of the SANDRA terminal for Collaborative Radio Resource Management (CRRM) and an approach to partition the functional entities between the IR and IMR for the configuration and reconfiguration of radio links during the connection management. The functional partitioning is based on two existing standards: the ETSI Broadband Satellite Multimedia (BSM) (ETSI, 2007) SI-SAP (Satellite Independent – Service Access Point) concept and the IEEE802.21 MIH framework (IEEE, 2009b). Both standards consider the separation of technology-independent upper layers from the technology-dependent lower layers to enable interoperability among heterogeneous networks. The BSM SI-SAP concentrates on the separation of upper layer functions from lower layer functions for satellite systems whereas the MIH framework defines media-independent functions for handover between heterogeneous mobile and wireless networks. The BSM SI-SAP concept for the interface between the IR and the IMR is extended to enable interoperability between satellite and terrestrial mobile wireless networks, while the IEEE 802.21 MIH framework is extended and adopted within the IMR to provide a uniform mechanism to control the different radio stacks.
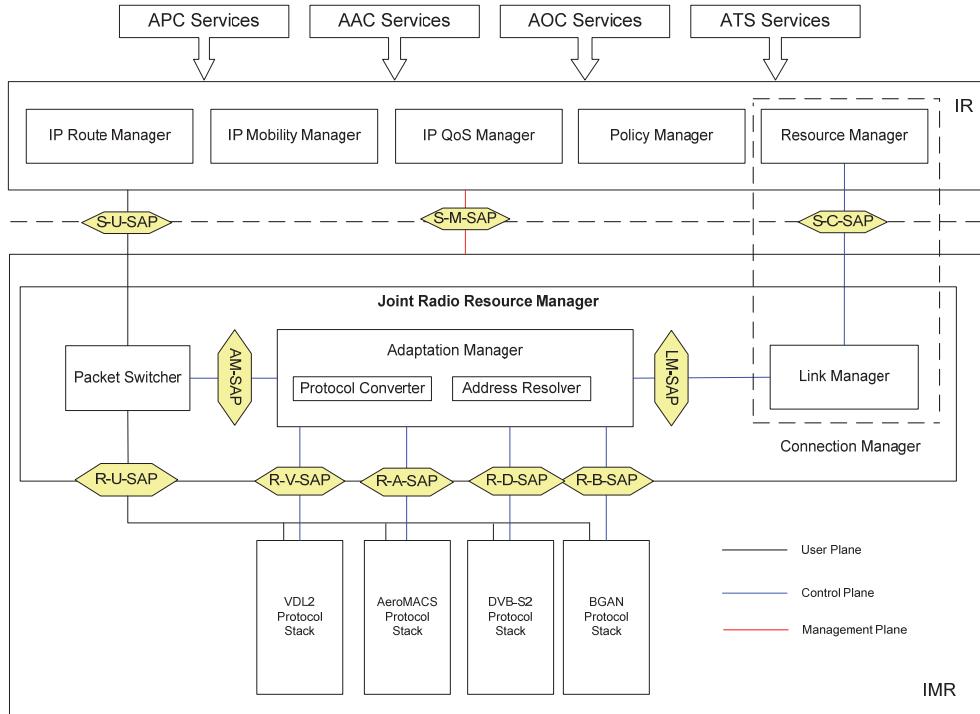
Fig. 6. CRRM architecture.

One key BSM SI-SAP parameter that is being utilised in SANDRA is the Queue Identifiers (QIDs) (ETSI, 2005) which are being used to identify queues in the IR. The QID is a 24 bit integer that permits a large number of queues to be defined for resource reservation and flow control functions. Combining MIH with the BSM QID and SI-SAP concepts will ensure QoS in the session management.

Fig. 6 shows the CRRM architecture. From the protocol stack point of view, the application layer supports APC, AAC, AOC and ATS services. The IR consists of the following RRM related functional entities: IP Route Manager, IP QoS Manager, IP Mobility Manager, Policy Manager and Resource Manager. The RM is the key functional entity in the IR to perform link selection in the CRRM mechanism. Detailed descriptions of these functional entities are out of scope of this chapter. The IMR consists of the JRRM, which is another key functional entity of the CRRM located in the IMR, and the radio protocol stacks for BGAN, DVB, AeroMACS and VDL2.

The JRRM, with its provision of adaptation functions and together with the link management functions, forms an Adaptation Layer between the IR and IMR for interfacing the network layer with the multiple radio stacks. This Adaptation Layer is primarily required for RRM purposes. It hides the underlying complexities due to multiple radio protocol stacks from the network layer and provides a uniform interface to control the multiple radios. The JRRM functional entities are further described below from the protocol stack perspective:

- Link Manager: Performs link selection and link configuration functions and together with the RM in the IR forms the Connection Manager (CM). The CM as a whole acts as a MIH user in the MIH framework.
- Adaptation Manager: Is primarily responsible for providing the various adaptation functions required for proper interfacing between the IR and the radio stacks. It consists of the following functions:
  - Protocol mapping: The request parameters like QoS parameters, security requirements, etc that are requested by the higher layers need to be mapped onto corresponding link-dependant parameters understood by the radio protocol stacks.
  - Address resolution: Different identities are used in the network layer and the radio protocol stacks to identify different connections. An address resolution table is maintained to provide correct mapping between the various identities used within the SANDRA system.
  - MIH Function (MIHF): The Adaptation manager also carries out a switching function equivalent to the MIH Function (MIHF) to support handover services including the Event Service (ES), Information Service (IS), and Command Service (CS), through service access points (SAPs) defined by the IEEE 802.21 MIH working group.
- Packet Switcher: Is responsible for switching data packets received from the IR in the user plane to the destined radio modules according to a packet switching table generated and passed by the address resolver in the Adaptation Manger (AM) during connection establishment. The packet switching table essentially contains the mapping of the QIDs onto the Link IDs. As a result, each data packet can be switched directly to the radio modules without passing through the AM in the user plane.

4.4.2 Collaborative RRM mechanism

The Collaborative RRM mechanism defined in SANDRA considers collaborative connection management, collaborative QoS management and collaborative mobility management in relation to admission control, packet scheduling and handover through cross-layer collaboration between functional entities in the IR and the IMR. While the IR is responsible for managing the network layer traffic flows, the IMR is responsible for the link layer connections. The collaborating entities, the RM and the LM, are grouped into a single cross-layer entity - the Connection Manager (CM). The relationships of the CM with the Policy Manager (PM), the IP QoS Manager (IQM), the IP Mobility Manger (IMM), IP Route Manger (IRM) and the AM as well as the mapping of IP Queues onto QIDs and Link IDs are depicted in Fig. 7.

The SANDRA Sessions mechanism is used for connection management and connection bindings between the IR and the IMR in order to provide efficient RRM and to meet the QoS requirements. Sessions are used to map corresponding data queues within the IP layer and the link layer. The term "session" corresponds to a single connection between a given IP queue and a link-layer queue. There is a one-to-one mapping between the IP queues and the sessions. In BSM, the Queue Identifier (QID) is used for mapping IP queues to BSM queues. This QID concept has been adopted and extended within SANDRA to identify the IP queues in the IR and its corresponding session with the IMR. Hence, each session is identified by a unique QID.
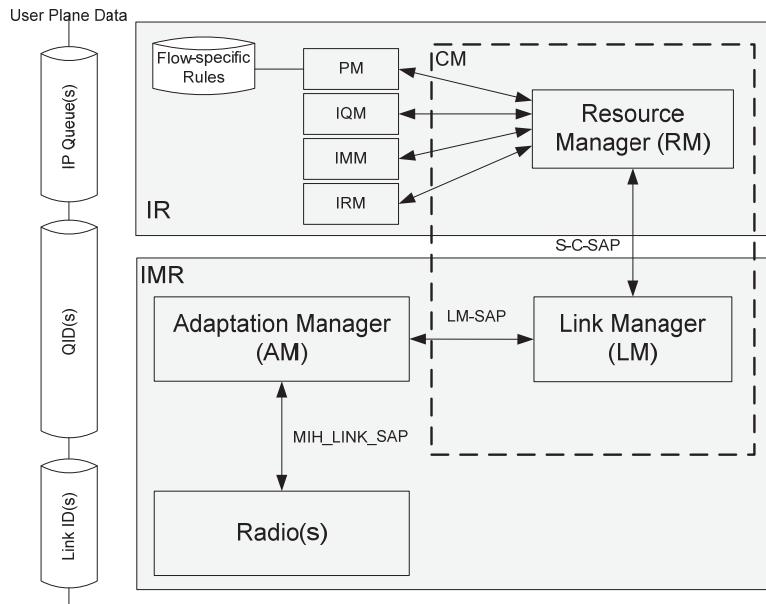
Fig. 7. Cross-layer collaborative connection management.

4.4.3 Collaborative connection management

Connection management functions including connection establishment, connection termination and connection modifications are carried out by the CM that enables decisions on network layer and link layer connections to be made by the RM in the IR and the LM in the IMR collaboratively. To establish a session for an application, the RM in the IR sends a RR to the LM, specifying the QoS parameters, user preferences, set up policy, etc. As the BSM SI-SAP interface is used between the IR and the IMR, this RR maps to the SI-C-Queue-OPEN-req message. The LM will configure the selected radio link to reserve bandwidth for the session according to the requirements specified in the RR. BSM QID is used between the IR and IMR to uniquely identify a session for the purpose of routing, scheduling, handover and session initiation/termination. Each session is mapped onto a radio specific connection identified by a Link ID used in different radio modules, such as a connection in AeroMACS or a PDP context in BGAN.

As shown in Fig. 7, the PM has knowledge on flow-specific rules and policies for individual applications and end-systems. These policies may govern the decision on which radio links can be used for different flows. Such policies may be based on the type of applications, type of planes, the location of the plane and flight path, etc. It may also be based on security and regulatory requirements that may restrict these flows to be transported over a specific radio link. Flows that do not have such strict requirements specified by the policy manager may be transported over a set of available links that satisfies the application QoS.

The Resource Manager in the IR may receive service requests from the applications in the higher layers specifying the QoS required by the applications. The resource manager will check with the PM to determine if there are any flow-specific rules already present for this particular applications that may be used for the suitable link selection. There are

applications that do not explicitly send a request to the IR specifying the required QoS but will start sending the application data directly to the IR. An example is the data sent from passenger specific application (e.g. web-browser). The Resource Manager will also be responsible for managing the resources required for such traffic. As a key part of the Collaborative RRM mechanism, the Resource Manager carries out the first level of decision making. It is responsible for deciding when new resources are required, when resources are released, etc. It will also perform link selection decision upon receiving a dedicated RR for a given application.

The Link Manager in the IMR is responsible for controlling the radio links and performs the second level of RRM related decision making for connection establishment. In the case of a general RR, the LM will select the most suitable link by mapping the application QoS requirements onto the resource availability and the quality of available links. The radio link that can most satisfy the QoS requirements will be selected and a general session between the IR and the selected radio link will be established.

4.4.4 Cross-layer collaborative QoS management

In relation to satisfying the QoS requirements upon a service request, the IQM in the IR will control and manage the IP Queues. On receiving data from the higher layers, the IP QoS manager performs packet classification based on the type of application and perform packet marking using Diffserv code points. Codes corresponding to the QoS requirements are added to the IP header of each packet before sending it to the IMR. The IQM in the IR also performs packet level scheduling of all incoming application packets based on their QoS requirements. The IR sees the different sessions between the IR and the IMR as different data pipes through which different data needs to be sent. Except under the situation when the IR submits a dedicated RR, data flow can be sent over any available sessions that may satisfy its QoS requirements.

The IMR needs to be able to also setup appropriate link-layer connections that meet the desired QoS that is requested by the IR. This requires mapping the higher layer QoS parameters to the link-specific QoS parameters. If the radio link cannot meet the desired QoS then another suitable link may be selected that could satisfy the QoS. If none of the available radio links is able to meet the desired QoS then the session request will either be accepted but with a degraded QoS or rejected if the minimum QoS cannot even be supported. In the latter case, the IR may then re-issue the resource request with the modified QoS parameters.

The IP QoS Manager in the IR is responsible for monitoring the IP queues to make sure that there are no packet drops within the system. The Packet Switcher in the IMR is also responsible to monitor any packet drops. These performance metrics need to be reported to the management Unit in the IR via the management plane. When the existing sessions are not able to satisfy the QoS needs of the application, then new session may be setup or additional resources may be requested on the existing radio links. This would require QoS re-negotiation with the ground networks.

4.4.5 Cross-layer mobility management

The SANDRA system supports multi-homing, where the IR can be connected to multiple ground networks via different radio links at any given time. Due to location constraints, handover support across different radios is required. For example, AeroMACS would primarily be available at the airports during taxiing, taking-off and landing whereas

satellites will be the primary means for communications when the airplanes are at cruising attitude. In addition, an airplane may move out of coverage of a given satellite link and may enter into another. The fast movement of the airplanes presents another complexity for mobility management in terms of handover.

In SANDRA, NEMO (Devarapalli et al., 2005) will be used by the IR for providing local and global mobility solutions and seamless mobility across the different networks. The IR and the IMR work in a collaborative manner to provide a cross layer mobility management solution. The IR may request the IMR to handover sessions from one radio link to another if there are some rules that dictate that different links may be used by an application during different phases of the flight. The IMR will also periodically monitor the link conditions and if it detects that a given link is no longer available then it will initiate different handover procedures based on the type of the associated sessions.

In the case of a general session, the Link Manager will select another suitable active link that satisfies the QoS requirements for this session. The Link Manager will then handover the session from the old link to the new link and informs the IR about the handover. The IR may then initiate the NEMO/Mobile IP signalling with the ground nodes.

In the case of a dedicated session, the LM will inform the IR about the change in link conditions associated with the dedicated session thereby triggering handover. The Resource Manager will perform suitable link selection for this session and inform the IMR of the newly selected link.

4.5 General RRM procedures

This section will introduce the general RRM procedures. An overview of these procedures is firstly given:

- Bootstrapping is the procedure of a radio module starts when the SANDRA terminal is turned on. It is very similar to the procedure of radio link up. A new QID is created at the same time and sent to the IR. It will be used by a new Resource Request.
- Connection establishment procedure is triggered either by a radio link detected or the IR which identifies it is required by the new application. This procedure will reserve resources on the particular radio links and setup the mapping between the IP queues to the radio links in order to transmit the data traffics from the user plane applications via the radio links.
- On some radio links, the provided services by the connections on them can also be modified. The connection modification procedure provides the IR with capability to modify the QoS profiles of the established connections, when it is necessary. If the modification is failed, two different procedures are performed based on the type of the established sessions:
  - For the dedicated session, the IMR directly reports to the IR about the failure of the connection modification.
  - For the general session, the IMR firstly tries to handover the session to the other radio link. If the handover procedure is successful, it will update the handover results with the IR; otherwise, it will reports to the IR about the failure of the connection modification.
- A radio link can become unavailable caused by any reasons. This triggers the radio link down procedure. When the IMR detects that a particular radio link is down, it will firstly identify all the established sessions on the radio link. For the dedicated sessions,

it will inform the IR that the sessions need to be disconnected. For the general sessions, it will try to handover them to other radio links firstly. If the handover for a general session is failed, the IMR will inform the IR that the session needs to be disconnected.

- Connection disconnect procedure provides the IR with the capability to terminate an ongoing session, when it identifies the session is not needed anymore.
- Handover is a very important RRM functions. There are two handover procedures provided in the SADRA system:
  - The IMR made handover decision for an ongoing session in order to enhance efficient RRM without effecting its QoS satisfaction in the lower layer. Normally this procedure is triggered when the IMR detects that a radio link is detected/down.
  - The IR made handover decision for an ongoing session in order to perform the mobility management in the upper layer.

The message sequence charts in Fig. 8 demonstrate how MIH primitives can incorporate BSM SI-SAP primitives for general session establishment (link selection by LM) and mobile controlled handover. The BSM SI-SAP primitives are shown as the signalling messages carried over the interface between the IR and IMR. These SI-SAP primitives will trigger a sequence of MIH link independent primitives, which will further trigger the link dependent primitives.

As seen in Fig. 8 (a), the resource request in a new session establishment procedure is handled by the ETSI BSM SI-C-Queue_Open-Req primitive that demands specific QoS requirements to be fulfilled by the IMR link setting. Upon reception of this primitive, the IMR makes use of MIH primitives to check the link status of each available radio technology then perform the link selection function to establish L2 connection on the selected radio technology. Finally, the ESTI BSM SI-C-Queue_Open-Cfm primitive is used by the IMR to confirm the establishment of L2 connection with the IR.

Fig. 8 (b) presents the layer 2 connection establishment procedure for handover using the ETSI BSM SI-C-Queue_Modify-Req primitive that indicates a new queue modify request due to the unavailability of resources on a given link or the detection of a newly available link that triggers a handover event. Consequently, QoS re-negotiation is required on the new link. This phase is then accomplished by making use of both ETSI BSM and MIH primitives as can be seen from the first three signalling message exchanges between the IR and the IMR.

4.6 Performance analysis of RRM procedures

To evaluate the performance of the RRM procedures, time delay analysis has been carried out based on the message sequence charts. The various wired and wireless links and interfaces between different network components shown in Fig. 8 have been considered. All messages involved in the procedure like connection establishment and handover have been taken into account in calculating the different delay components. In general, the total time taken to transmit a single message over any given link, D_{Total} can be expressed as the sum of four delay components (Pillai & Hu, 2009):

$$D_{Total} = D_{Prop} + D_{Proc} + D_{Trans} + D_{queue} \qquad (1)$$

Where, D_{Prop} is the propagation delay, D_{Proc} is the processing delay, D_{queue} is queuing delay and D_{Trans} is the transmission delay. The general queuing delay D_{queue} for any network entity,
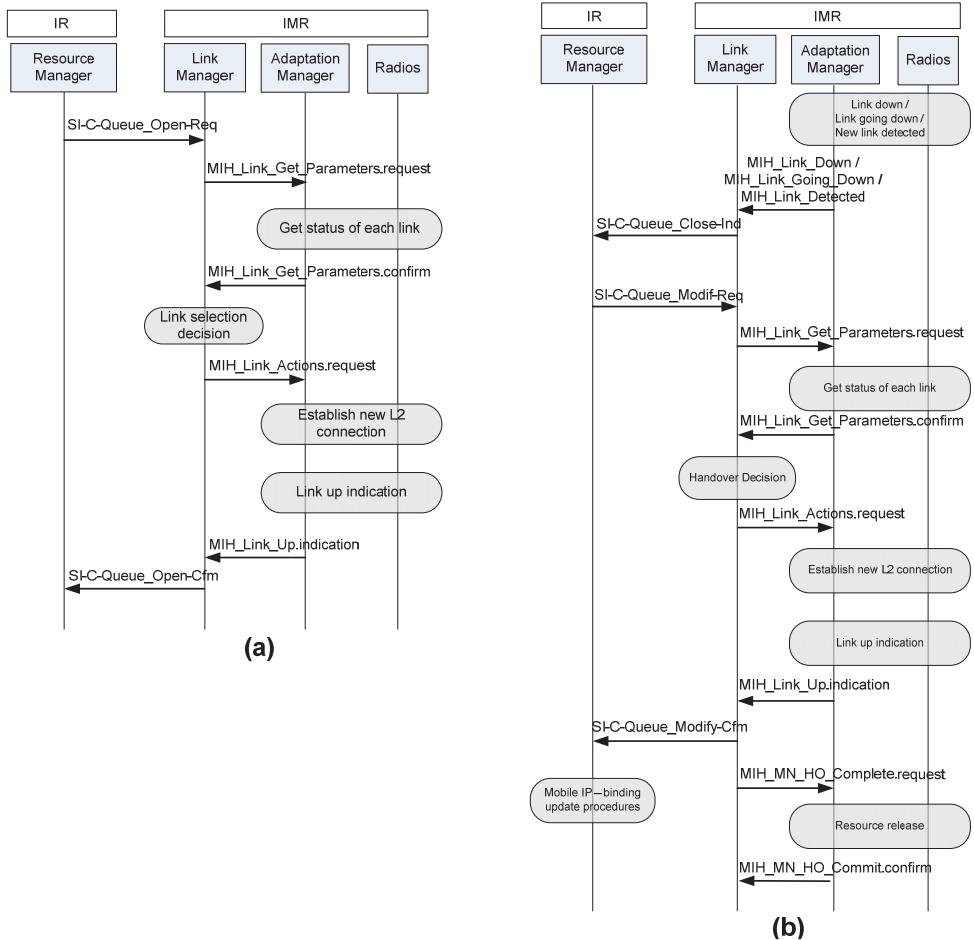
Fig. 8. (a) General session establishment and (b) mobile controlled handover.

based on an M/M/1 queuing model can be expressed as $D_{queue} = \dfrac{1}{(\mu C - \lambda)}$ where μ is the service rate, C is channel capacity and λ is the arrival rate.

Table 1 presents the various parameters adopted for evaluating the total delay for the two procedures. The total delay for the session establishment procedure is expressed as follows (Ali et al., 2011):

$$D_{Total} = P(D_{Proc}) + 2\left\{ D_{Prop} + D_{Trans} + D_{Proc} + \left(\frac{1}{\mu C_i - \lambda_{wired}} \right) \right\}$$

$$+ (K_{sel})\left\{ D_{Prop} + D_{Trans} + D_{Proc} + \left(\frac{1}{\mu C_j - \lambda_{wireless}} \right) \right\} \qquad (2)$$

In Equation 2, P represents the total number of messages exchanged between entities within the IMR where K_{sel} denotes the number of messages required for session establishment over the selected wireless link. C_i and C_j denote the capacity of wired and wireless communication channel.

Similarly, the signalling delay for the handover procedure shown in Fig. 8 (b) is given by equation (3), where Q represents the total number of messages exchanged within IMR entities.

$$D_{Total} = Q(D_{Proc}) + 3\left\{D_{Prop} + D_{Proc} + D_{Trans} + \left(\frac{1}{\mu C_i - \lambda_{wired}}\right)\right\}$$

$$+\left[(K_{Sel})\left\{D_{Prop} + D_{Trans} + D_{Proc} + \left(\frac{1}{\mu C_j - \lambda_{wireless}}\right)\right\}\right] \tag{3}$$

Fig. 9 shows the total signalling delay during session establishment and during seamless vertical handover respectively. It can be seen from both figures that an increase in the arrival rate will cause an increase in the total signalling delay as a result of an increase in the queuing delay D_{queue}. Fig. 9 (a) illustrates AeroMACS exhibits the lowest delay for session establishment as its propagation delay is small and data rate is high. DVB-S2 has higher data rate than AeroMACS but incorporates high propagation delays. BGAN has the lowest data rate of 492 kbps and high propagation delays therefore it exhibits the highest total delay values in the graphs. The graphs also show that high data rate provides better results for high arrival rate. For example, the total delay for AeroMACS session establishment becomes more than that for DVB-S2 when the arrival rate goes beyond around 82 packets/sec. Similarly the handover delays for different handovers are shown in Fig. 9 (b). It is shown that handover delays from DVB-S2 to BGAN and AeroMACS to BGAN are the highest (nearly 2 seconds) which is due to large signalling overhead for handover and propagation delay in the BGAN network. The handover delay for handover, from DVB-S2 to AeroMACS is the lowest as target network (AeroMACS) has lowest propagation delay and low signalling over head for handover.
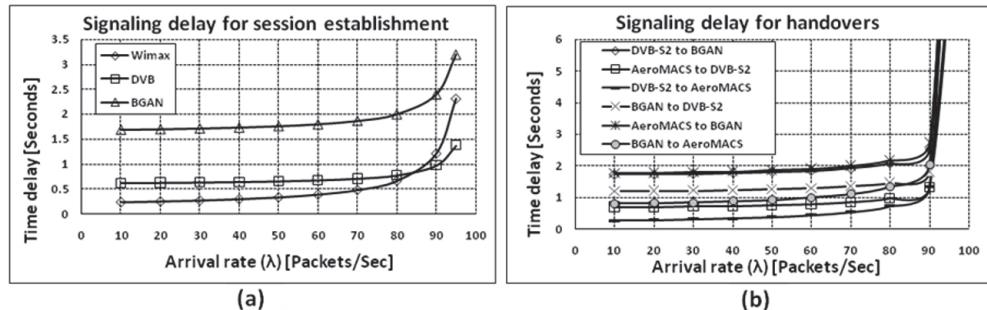

Fig. 9. (a) Signalling delay for new session establishment on different technologies and (b) Signalling delay to handover to different technologies.

No.	Parameter Name	Symbol	Value
1	Arrival rate at wireless interface	$\lambda_{wireless}$	10 to 95 s-1
2	Arrival rate at wired interface	λ_{wired}	10 to 95 s-1
3	Service rate	μ	100 s-1
4	Wireless channel capacity	C_j	45, 0.492, 80 Mbps
5	Wired channel capacity	C_i	100 Mbps
6	Wireless link Propagation delay	$D_{Prop.}$	250ms, 167µs
7	Wired link propagation delay	$D_{Prop.}$	5ns
8	Wireless transmission delay	$D_{Trans.}$	293µs, 2.4µs, 23µs
9	Wired transmission delay	$D_{Trans.}$	4.4 µs
11	Average packet processing delay	$D_{Proc.}$	5ms
12	Number of messages exchange required between mobile node and network entity for session setup.	$K_{sel.}$	6(BGAN), 9(AeroMACS), Approx. 2(for DVB-S2 receiver synchronization)
13	Average packet length	BGAN	18 bytes
		DVB-S2	24 bytes
		AeroMACS	129 bytes
		Wired link	55 bytes

Table 1. Parameter value chart.

5. Conclusion

This chapter firstly gives a brief overview on the radio transport technologies adopted by the aeronautical communication system. Two existing standards: BSM concept and IEEE 802.21 MIH framework have been reviewed to enable interoperability among heterogeneous networks by considering the separation of technology independent upper layers from the technology dependent lower layers. To enable a close collaboration between the IR and the IMR, which manage the terminal's upper and lower layer functionalities respectively, for efficient RRM, a Collaborative RRM (CRRM) scheme has been proposed. A detailed description of the SANDRA network architecture and the CRRM functional architecture are presented to illustrate the seamless interoperability across the heterogeneous networks. The CRRM mechanism highlights the mechanisms and advantages of adopting ETSI BSM SI-SAP concept and the IEEE 802.21 MIH framework and splits the CRRM functions between the upper layers (layer 3 and above) and the lower layers (link layer and physical layer) of an aircraft terminal. A joint radio resource manager (JRRM) provides the abstraction layer between the IR and IMR for mapping higher layer functions into lower layer functions to enable collaboration. Through the CRRM scheme, the IR and IMR maintain a close collaboration to perform connection establishment functions and to support seamless handovers between different radio technologies.

To continue with the design of the CRRM scheme, the behaviours of the mechanism and collaborative RRM procedures are given in this chapter. Two detailed general message sequence charts have been provided to demonstrate the combined use of MIH and extended ETSI BSM primitives for general RRM procedures like session establishment and handover management. Finally an analytical model is used to measure the signalling delay for the RRM procedures. The results show that DVB-S2 offers more bandwidth and is more tolerant to an increase in arrival traffic. BGAN having lowest data rate and high propagation delay exhibits the highest total delays. AeroMACS, which will be used when an aircraft

approaches the airport, having low propagation delay and high data rate, shows the lowest total delay. Since DVB-S2 has the same propagation delay as BGAN but with a higher data rate, its delay performance is better than AeroMACS under high arrival rate.

6. Acknowledgment

7. References

Ali, M., Xu, K., Pillai, P., &Hu, Y.F. (2011). Common RRM in Satellite-Terrestrial Based Aeronautical Communication Networks, PSATS 2011, Spain, February 2011.

ARINC. (2011). Aircraft Communications Addressing and Reporting System (ACARS), Available from http://www.arinc.com/products/voice_data_comm/acars.html.

Denos, R. (2010). *Aeronautics and Air Transport Research - 7th Framework Programme 2007-2013 - Project Synopses*, Volume 1 Calls 2007 & 2008, European Commission.

Devarapalli, V., Wakikawa, R., Petrescu, A., & Thubert, P. (2005). *Network Mobility (NEMO) Basic Support Protocol.* RFC 3963.

ESA. (2003). BGAN Project Objectives, Available from http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=11366

ETSI. (2005). *Technical Specification, Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) Common air interface specification; Satellite Independent Service Access Point (SI-SAP).*TS 102 357 V1.1.1.

ETSI. (2007).*Technical Report, Satellite Earth Station and Systems (SES); Broadband Satellite Multimedia (BSM); Services and architectures.*TR 101 984 V1.2.1.

ETSI. (2009a).*Technical Report, Reconfigurable Radio System (RRS); Software Defined Radio Reference Architecture for Mobile Device.* TR 102 680 V1.1.1.

ETSI. (2009b).*Digital Video broadcasting (DVB) 2nd generation framing structure, channel coding & modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications (DVB-S2).*EN 302 307 V1.2.1.

ETSI. (2009c). *Technical Report, Reconfigurable Radio System (RRS); Radio Base Station (RBS) Software Defined Radio (SDR) status, implementations and costs aspects, including future possibilities.* TR 102 681 V1.1.1.

EUROCONTROL. (2001). FAA/EUROCONTROL Memorandum of Co-operation, Available from http://www.eurocontrol.int/moc-faa-euro/public/subsite_homepage/homepage.html

EUROCONTROL. (2006). *Long-Term Forecast, Flight Movements (2006 - 2025) v1.0,* EUROCONTROL.

EUROCONTROL/FAA. (2007). *Action Plan 17: Final Conclusions and Recommendations Report*, EUROCONTROL/FAA Memorandum of Cooperation.

EUROCONTROL. (2009). *IEEE 802.16e System Profile Analysis for FCI's Airport Surface Operation*, EUROPEAN AIR TRAFFIC MANAGEMENT

Fantacci, R., Marabissi, D., & Tarchi, D. (2009). Adaptive Scheduling Algorithms for Multimedia Traffic in Wireless OFDMA Systems. *Physical Communication*, vol. 2, pp. 228-234.

Giambene, G. (2007).*Resource Management in Satellite Networks: Optimization and Cross-Layer Design, 1st ed*, Springer.

Homans, A. (2002). The Evolving Role of the Communication Service Provider. Integrated CNS Technologies Conference and Workshop, May 2002.

ICAO .(2001). *Manual on VHF Digital Link (VDL) Mode 2*, Doc 9776 AN/970.

IEEE. (2009a). *Part 16: Air Interface for Broadband Wireless Access Systems.* IEEE Std. 802.16.

IEEE. (2009b). *Local and Metropolitan Area Networks - Media Independent Handover Services.* IEEE Std. 802.21.

INMARSAT. (2003). *INMARSAT BGAN System Definition Manual*.

INMARSAT. (2011). BGAN, Available from http://www.inmarsat.com /Services/Land/Services/High_speed_data/default.aspx

Jilg, G. (2002). *INMARSAT - products and strategies*, Workshop on Satellites in IP and Multimedia, Geneva.

Kumar, G. S. A., Manimaran, G., & Wang, Z. (2009). Energy-Aware Scheduling with Probabilistic Deadline Constraints in Wireless Networks. *Ad Hoc Networks*, vol. 7, pp. 1400-1413.

Kuroda, M., Saito, Y., Ishizu, K., & Komiya, R. (2006). *Clarification of MIH_NMS_SAP*, DCN: 21-06-0786-00-0000.

NASA. (2005). *Technology assessment for the future aeronautical communications systems*, NASA ITT Industries, NASA-CR-20050213587

Pillai, P., & Hu, Y. F. (2009). Performance analysis of EAP methods used as GDOI Phase 1 for IP multicast on Airplanes, WAINA'09 international Conference, June 2009.

SANDRA. (2011). SANDRA Concept, Available from http://www.sandra.aero

Design Aspects of a Testbed for an IPv6-Based Future Network for Aeronautical Safety and Non-Safety Communication

Oliver Lücke and Eriza Hafid Fazli
TriaGnoSys GmbH
Germany

1. Introduction

In this chapter, the development of a networking testbed to test and validate IPv6-based protocols for future Air Traffic Management (ATM) network is presented.

The development was originally initiated within the EC FP6 project NEWSKY (Networking the Sky for Aeronautical Communications), which aimed at developing a concept for a global, heterogeneous communication network for aeronautical communications, based on IPv6 protocol stack. The NEWSKY network integrates different applications (ATS, AOC, AAC, and APC) and different data link technologies (legacy and future long range terrestrial radio, satellite, airport data link, etc.) using a common IPv6 network layer. For proof-of-concept, the NEWSKY testbed has successfully implemented NEWSKY network mobility, handover, and quality of service solutions, and tested and demonstrated them over real satellite link (Fazli et al., 2009).

Also the EC FP7 project SANDRA (Seamless Aeronautical Networking through integration of Data-Links, Radios and Antennas) aims at designing and implementing an integrated aeronautical communication system and validating it through a testbed and, further, in-flight trials on an A320 (SANDRA web page, 2011). Central design paradigm is the improvement of efficiency and cost-effectiveness by ensuring a high degree of flexibility, scalability, modularity and re-configurability.

Whereas the NEWSKY testbed is considered to be a *proof-of-concept*, the SANDRA testbed will represent a *proof-of-principle prototype* aircraft communication system, integrating prototypes developed and implemented in SANDRA, comprising AeroMACS, Integrated Modular Radio (IMR), Integrated Router (IR), and a novel Ku-band electronically steerable antenna array.

SANDRA focuses on the air-to-ground communication and on the development of the on-board airborne SANDRA terminal. The SANDRA terminal, in particular the IR and the IMR have to jointly implement the capabilities of resource allocation among heterogeneous link access technologies and link reconfigurability (e.g. when new links become available or previously available become unavailable, including handover between links). The required technology-dependent functions (such as control of the heterogeneous link technologies) reside in the IMR, whereas technology-independent functions are implemented in the IR, while using IP to achieve convergence and interoperability between the different link access

technologies. Although the focus is on the airborne terminal, the testbed of course also has to implement a ground network side.

The main aspects of the testbed design and the testbed core concepts and components are presented in this chapter, namely the security and QoS (quality of service) provision concepts for segregation of safety and non-safety domains in an integrated system, the Integrated Router, IPv6 and IPv4 internetworking.

This chapter is organised as follows.

First, in Sec. 2 the NEWSKY project will be shortly described, including the study objectives and the architecture of the IPv6 network testbed for NEWSKY demonstration is also described.

The following Sec. 3 contains a short overview of the SANDRA main goals and testbed purpose. The main differences of the SANDRA testbed from the NEWSKY testbed will become evident. This includes improvements in the protocols and configuration taking into account lessons learnt in NEWSKY, additional features such as automatic modem and bandwidth control, Integrated Modular Radio (IMR), and the additional requirement that the testbed shall be integrated in an Airbus A320 for in-flight test.

Sec. 4 focuses on presenting the overall SANDRA testbed network architecture. This includes a reference network architecture, which is independent of restrictions specific to the testbed implementation, and the actual testbed network architecture which has to take into account various constraints, e.g. unavailability of IPv6 satellite access networks.

Sec. 5 presents in some more detail selected topics related to the SANDRA testbed which are specifically interesting aspects of implementation and investigation, such as IPv6 over IPv4 network traversal and header compression.

Finally, Sec. 6 gives a short summary and presents the next steps and schedule for the testbed implementation and laboratory and flight trials.

2. NEWSKY testbed overview

The main objective of the laboratory testbed development in NEWSKY was to implement some basic functionalities of the NEWSKY networking design, in particular with respect to mobility, security, and QoS aspects. However, due to the constraints within the project, not all design aspects were implemented in the testbed. This section describes the NEWSKY testbed architecture, components, and points out the differences between the design and the implementation wherever applicable.

The laboratory testbed consists of two main subnetworks: the *airborne mobile network* representing an aircraft, and the *ground network* representing a connected ATC and airline network, and the public Internet. The two main subnetworks are connected by two different data links. The network architecture is depicted in Fig. 1.

The airborne network consists of the Mobile Network Node (MNN) and the Mobile Router (MR). The MNN represents the airborne user terminal, which could belong either to the cockpit (e.g. an interface for pilot-ATC communication) or to the cabin (e.g. passenger's laptop). As part of the NEMO (Network Mobility) protocol (see description below), a Home Agent (HA) is installed in the ground network. The Correspondent Node (CN) acts as the other end point of the communication, e.g. it could be the air traffic controller, or the airline. An interface to the public Internet is also implemented to emulate passenger Internet connectivity. Further testbed components are described below.

Improvement of the testbed was continued after NEWSKY was completed. Some of the improvements include the integration of Multiple Care-of Address (MCoA) extension for

NEMO (Wakikawa et al., 2009), and the implementation of NEWSKY IPsec based domain separation. The testbed served as an initial proof-of-concept of the NEWSKY's future ATM networking design, and is to be further developed into a prototype within the SANDRA project.

2.1 Network mobility protocol

Mobile IPv6 (MIPv6)/NEMO (Johnson et al., 2004; Devrapalli et al., 2005; Perkins et al., 2011) and its extensions have been selected by the International Civil Aviation Organization (ICAO) Aeronautical Communications Panel Working Group I (ACP WG-I) to be the solution for global network mobility in future IPv6-based aeronautical telecommunication network (ATN). NEWSKY took the same approach by specifying MIPv6 as its solution for mobility. Network Mobility (NEMO) protocol is introduced to extend MIPv6, enabling the mobility of a complete network instead of just a single host. The testbed uses Linux based NEMO protocol from the Nautilus6 project (Nautilus web page, 2009).

2.2 Data links

Two data link technologies are integrated in the testbed, namely the Broadband Global Area Network (BGAN) system from Inmarsat, and an emulated terrestrial link of a potential future air-ground data link technology, L-DACS-1 (L-band digital aeronautical communications system). The satellite link is a real physical link through the Inmarsat I4 satellites, whereas the emulator is software based, introducing transmission delay and limiting the data rate of an Ethernet link, representing the physical and data-link layer of L-DACS-1.
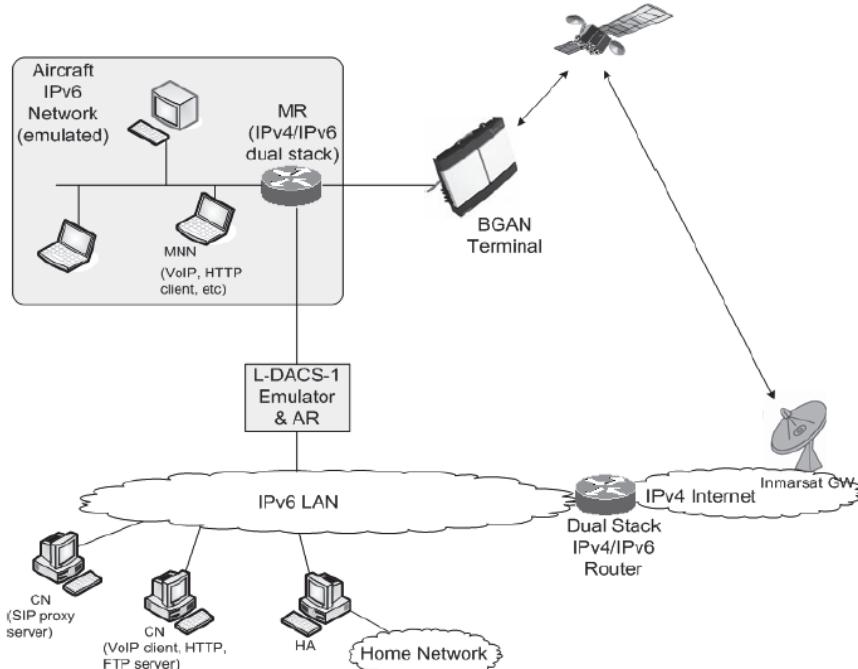

Fig. 1. NEWSKY laboratory testbed network architecture.

2.3 IPv4 network traversal

Although NEWSKY focuses only on IPv6, in reality the Inmarsat BGAN system is based on IPv4. In the testbed, we implemented the Network Crossing via Translation protocol (NeXT) to efficiently traverse the satellite IPv4 network. NeXT is an overhead optimised NAPT-PT (Network Address and Port Translation - Protocol Translation) mechanism (Via et al., 2009).

2.4 Applications

Future ATM communications will rely more on data services, and the usage of voice as in the current VHF-based radio system will be greatly reduced. Therefore in the testbed we implemented some representative IP-based applications, e.g. VoIP and FTP data transfer.

2.5 QoS management

NEWSKY specifies DiffServ to guarantee the prioritisation of message coming from different applications. In the testbed this is implemented using Linux based *tc* tool to direct IPv6 packet flows into different queue classes, instead of putting traffic class marks directly to the packet as in DiffServ.

2.6 Domain separation

For safety and security reasons, the on-board aircraft mobile network is divided into safety-critical subnetwork for ATS/AOC applications, and non-safety-critical subnetwork for APC applications. In the current airborne network this separation is achieved physically, i.e. separate networks for safety and non-safety domains are deployed. NEWSKY proposed a future airborne network architecture, where the separation is achieved logically using IPsec security tunnels. At the time of the project, however, this was not implemented, and a firewall is configured at the MR to enable a simpler separation of the domains.

3. SANDRA key elements

A key contribution of the SANDRA project to the definition of future aeronautical communication system is the integration on different system levels to overcome the drawbacks (such as limited coverage of direct air-to-ground data links, high delay of satellite systems, etc.) of existing aeronautical communications systems based on single radio technologies and individual radio access systems:

- Integration of services from the safety and non-safety domains (ATS, AOC, AAC, APC).
- Integration of different networks, based on interworking of different radio access technologies through a common IP-based aeronautical network whilst maintaining support for existing network technologies (ACARS, ATN/OSI, ATN/IPS, IPv4, IPv6).
- Integration of different radio technologies in an Integrated Modular Radio (IMR), applying concepts for Integrated Modular Avionics (IMA) to communication avionics.
- Integration of Ku- and L-band antennas in a single hybrid Ku/L band satellite communication antenna providing an asymmetric broadband link.
- Design and implementation of AeroMACS (Aeronautical Mobile Airport Communications System, based on mobile WiMAX standard 802.16e) for integrated multi-domain airport connectivity in C-band.

To prove all of the concepts developed in the SANDRA global system architecture the IMR will be capable of interfacing with a sufficient number of bearers, comprising Inmarsat

SwiftBroadband (SBB) class 6/7, Ku DVB-S2 (2nd generation Digital video broadcast, receive only), AeroMACS, VHF Voice and VDL Mode 2 (VDL2).

Closely related to the key contributions, the key requirements for SANDRA comprise

- Support of data and digital voice services for ATS, AOC, AAC and APC.
- Support of interfacing with legacy and future data links.
- Ensuring network interoperability comprising ACARS, ATN/OSI, IPv4, IPv6.
- Deploying IPv6 as unification.
- Compatibility with ATN/IPS.
- Performing link selection based on information from applications, policies, and link layer information.

A further key element of the SANDRA project is the testbed, which has the purpose to implement and demonstrate as many of these SANDRA key elements as possible for validation of the overall SANDRA concept and architecture up to Technology Readiness Level (TRL) 5-6[1].

The SANDRA testbed design and implementation will also take into account more specific and comprehensive requirements definitions prepared within SANDRA in various Work Packages (WP). SANDRA activities being relevant for the testbed specification are "WP3.1 - Detailed network requirements", "WP3.2 - Network architecture", "WP3.5 - IPS network design", "WP4.2 - IMR interface", "WP7.1 - identification of specific hardware available for the testbed", "WP7.2 - Use cases and scenarios to be covered by the testbed", and "WP7.3 - Identification and Implementation of Applications".

According to the envisaged TRL, a central goal of the SANDRA testbed is the development of a SANDRA *proof-of-principle prototype* (in contrast to NEWSKY's testbed for *proof-of-concept*), integrating the prototypes implemented in SANDRA (Integrated Router (IR), Integrated Modular Radio (IMR), novel Ku-band antenna, and AeroMACS) for validation&verification and also demonstration

    a. in a laboratory set-up of a high-fidelity representation of the target SANDRA system, including the airborne and the ground networks, and also comprising a realistic system environment for the airborne network (i.e. an aircraft) and

    b. during flight trials on an Airbus A320 (planned for 2nd quarter 2013).

In addition to the SANDRA prototypes all network elements will be implemented and integrated in the testbed that are necessary to create the high-fidelity representation of the SANDRA target system (e.g., global geographically distributed networks, mobility via NEMO including extensions, security provisions, techniques for improving efficiency (e.g. Robust Header Compression/RoHC, etc.), including the 4 domains (ATS, AOC, AAC, and APC) for the end systems (ES) which will host the applications, and all other entities in the airborne and in the ground networks (cf. Sec. 4 and 5 for details).

The detailed testbed requirements specification and implementation phase was started in Oct. 2010 and is expected to be on-going until end of 2nd quarter 2012. The subsequent integration of all components and testbed trials will continue until 1st quarter 2013.

In addition to the above stated main purpose of the testbed, further suggestions and new algorithms from SANDRA activities running in parallel to the testbed design and implementation will be considered for implementation in the testbed.

---

[1] TRL 5: components and/or breadboard validation in *relevant* environment; thus, flight trials will be performed within SANDRA. TRL 6: system & subsystem model or prototype demonstration in a relevant environment (ground or space). The highest TRL (TRL 9) corresponds to the actual system "flight proven" through successful mission operations.

Thus this chapter shows work in progress and cannot provide all details of the SANDRA testbed, in particular presentation of results will have to wait until the testbed implementation is finalised and trials have been performed.

Finally, as stated above, the SANDRA testbed will be deployed in a laboratory environment and in an aircraft for flight trials. The testbed set-up for the flight trials will be subject to some restrictions, e.g. no Ku-band link will be available for the flight trials because of the too high costs for the installation and certification of the Ku-band antenna on the aircraft fuselage. However, we will not further address in the following the differences between the two testbed set-ups (laboratory and flight trials) and will not address in more detail the restrictions for the flight trials.

4. SANDRA testbed network architecture overview

4.1 Testbed reference network architecture

In Oct. 2010, the SANDRA testbed network architecture was finalised in a Software Requirements Document (SRD). In this SRD, the SANDRA testbed reference network architecture is defined, which initially is independent of the restrictions specific to the testbed (e.g. no IPv6 satellite network); a schematical overview is shown in Fig. 2.[2] ATC (air traffic control, safety related) ground networks and (public) Internet (non-safety related) and Correspondent Nodes (CN) for the 4 domains are also located in other geographical regions (region 2 and 3 in reference architecture diagram) but not shown to keep the figure simpler.

The network nodes in the airborne network are:

- Airborne mobile network nodes (MNN) in the domains ATS, AOC, AAC, and APC; these are hosting the applications. Note that middleware and/or transition gateways may need to be deployed to integrate end-nodes that do not provide IPv6 interfaces.
- Firewalls at the entry/exits points of the 4 operational and non-operational domains to inhibit entry/exit of packets from malicious hosts (intentional or unintentional). Firewalls will not be implemented in the testbed because no penetration or performance tests will be performed.
- IPsec (IPv6) gateways to provide authentication and integrity, with or without encryption, for safety (ATS/AOC) and non-safety related data traffic (AAC/APC). Further, IPsec enforces in the airborne network segregation of the safety domains from the non-safety domains. Also VLANs, port or tag based, are possible options to further enforce segregation but, like the firewalls, will not be implemented in the testbed.
- The Mobile Router (MR, IPv4 and IPv6) (also referred to as the Integrated Router, IR), is responsible for controlling the modems (for the fall-back solution[3]) or the IMR, resource allocation, QoS provision at IP packet level, etc. The NEMO related functionalities of the

---

[2] Because NEMO is used for mobility support, the reference architecture uses the nomenclature of (Devarapalli, 2005), comprising Mobile Network Nodes (MNN), Mobile Router (MR), Home Agent (HA), and Correspondent Node (CN).

[3] Already included in the SANDRA proposal is a fall-back solution *without* the IMR, in which case the IR directly interfaces with respective COTS (Commercial off-the-shelf) modems (SBB, DVB-S2, WiMAX). The fall-back solution is included to mitigate a potential risk of delay in the IMR implementation due to the high complexity of this prototype development. In this case, the IR has to implement some functionalities of the IMR, such as modem control and handover management.

MR are specified in (Devarapalli et al., 2005) and (Wakikawa et al., 2009), including, e.g., registration of multiple CoAs at the HA.
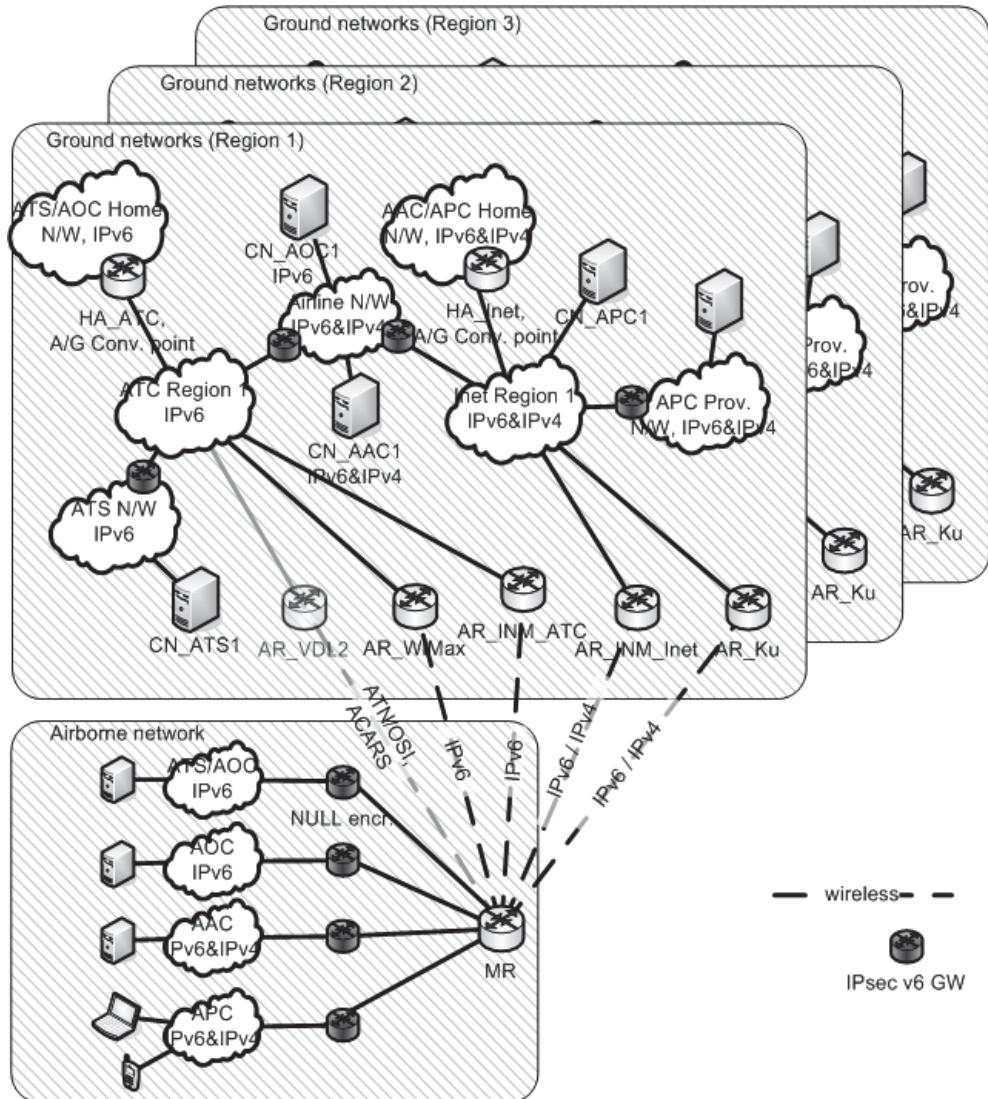

Fig. 2. SANDRA reference network architecture.

The network nodes in the ground network are:
- Access routers (AR); at least one AR is deployed for each access technology (different provider may deploy different ARs or several geographically distributed ARs are deployed together with multiple radio access stations). Access routers provide CoA to the MR.

- Home Agents (HA); at least one HA is deployed. The functionality of the HA are again specified in (Devarapalli et al., 2005) and (Wakikawa et al., 2009). An HA further represents an air-ground routing convergence point, being aware of all routing paths (multiple CoAs) to the aircraft MR and forwarding traffic by policy routing to the MR on a particular routing path, e.g. based on QoS requirements.
- IPsec gateways being the counterparts of the airborne IPsec gateways securing the ATS, AOC, AAC, and APC domains.
- Correspondent Nodes (CN) in the domains ATS, AOC, AAC, APC; these are hosting the respective applications.

Depending on geographical location, the MR connects to a respective AR. ATC regional ground subnetworks are interconnected to implement global connectivity and the same holds for the Internet and its regional subnetworks.

The ES (MNN and CN) are not addressed in detail in this chapter, which only deals with the IP network up to the interfaces where the ES are connected to.

Although integration of legacy data links and interoperability with ACARS, ATN/OSI, IPv4 are considered essential for SANDRA (cf. Sec. 3), in the testbed the focus is on IP networking (IPv6 and IPv4). However, integration of legacy ES and applications (ACARS, ATN/OSI, analog voice) and VDL2 as access network into the testbed is currently still under investigation in terms of the details of implementation and thus not included here.

4.2 Testbed network architecture

In contrast to the reference SANDRA testbed network architecture presented in the previous section, the actual testbed architecture and implementation will have to take into account various constraints, such as the necessity to include IPv4 access and ground networks also for ATS/AOC because of the current unavailability of IPv6 satellite networks.

The testbed ground-network will mainly be set-up at TriaGnoSys (TGS) premises in Oberpfaffenhofen, Germany. Also the airborne side will initially be set-up at TGS premises as well.

Due to limitations for the flight trials that do not apply to the laboratory set-up (e.g. no Ku-band antenna on aircraft), only a subset of the laboratory equipment of the airborne network will be used in the flight trials and installed in the aircraft.

Note that, as already stated above, the actual testbed network architecture will comprise IPv4 access and ground networks between the IPv6 only ATS/AOC airborne and ground networks due to the lack of support of IPv6 in the satellite networks available for the testbed implementation.

In contrast to ATS/AOC, for AAC/APC the IPv4 access and ground networks are a desired element of the testbed, because support of IPv4 hosts and access networks is required for AAC/APC.

In an aeronautical communication system that is deployed world-wide, MNN and MR, AR, HA, and CN are generally geographically distributed. Latencies between nodes that are close to each other are typically small, whereas latencies between nodes that are far away from each other may be significantly larger (Ayaz et al., 2009). To reflect this, the testbed ground network is separated in three regions (three regions are considered sufficient for most scenarios, e.g. that AR, HA, and CN are located in at most three different regions); latencies between nodes in the same region are small, whereas latencies between nodes in different regions may be significantly larger.

Design Aspects of a Testbed for an IPv6-Based Future
Network for Aeronautical Safety and Non-Safety Communication
179

Further, it is currently not foreseen to implement separate ATC and public Internet ground networks (cf. reference network architecture from previous section). However, if later definition of use cases, scenarios, and applications, not known currently, will require this, then separate ground networks will be implemented by deployment of additional routers (possibly virtual).

The resulting SANDRA testbed network architecture is shown in Fig. 3.

As for NEWSKY, the mobility solution for the SANDRA testbed is NEMO (Devarapalli et al., 2005), including the possibility to register multiple Care-of-Addresses (Wakikawa et al., 2009). NEMO Route Optimisation (RO) will be considered (e.g. Global HAHA, Correspondent Router, etc.), however, implementation in the testbed will depend on an assessment of benefit and effort of the implementation, in particular if open source implementations are not yet available.

Use of multiple CoAs requires synchronisation of policy routing rules at the MR and the HA. This is addressed in the recent IETF RFC 6089 (Tsirtsis et al., 2011), which also will be considered in the SANDRA testbed.

Finally, NEMO Basic Mobility Support as specified in (Devarapalli et al., 2005) does not foresee IPv4 access networks. Two options allowing deployment of NEMO over IPv4 access networks will be considered for implementation in the testbed; first, RFC 5555 (Soliman, 2009) and, second, an efficient NAPT-PT based solution (Via et al., 2009) that was already successfully deployed in the NEWSKY testbed.
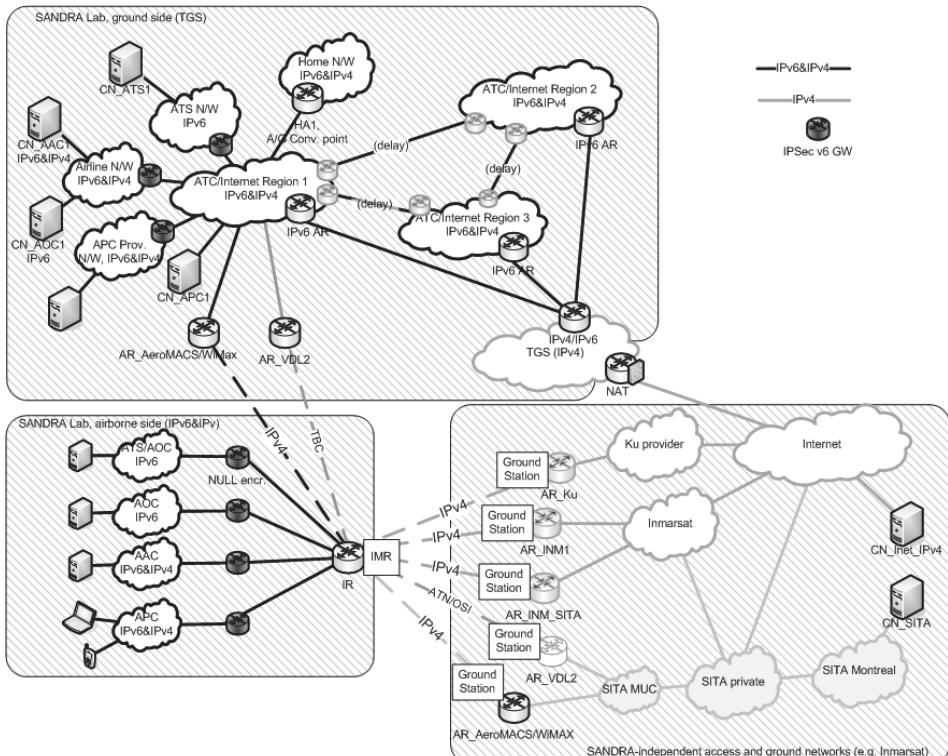

Fig. 3. Testbed network architecture overview, including IPv4 access and ground networks.

5. Selected topics in the SANDRA testbed

In the following, we highlight a selection of topics addressed in the SANDRA testbed.

5.1 IPv6 over IPv4 network traversal

For each connected access network, the IR obtains from the IMR (or directly from a modem in case of the fall-back solution) one or more IPv6 or IPv4 CoAs (e.g. SBB allocates an IPv4 address for each primary packet data protocol (PDP) context), which are provided by the respective AR (by which protocol this is specifically done may depend on the IMR interface to the IR, the specific modem, the access network etc., which are not yet defined in the necessary detail).

In case RFC 5555 is deployed for IPv4 network traversal, IPv4 CoA are registered at the HA via a binding update (Soliman, 2009). IPv6 packets, including NEMO signalling, are tunnelled between MR and HA in IPv4 and a UDP header is added in case that NAT (Network Address Translation) is present on the path.

When NeXT is deployed, traversal of the IPv4 access network for IPv6 packets is achieved via IPv4/IPv6 translation. A context is set-up between NeXT master (in the IR) and slave (e.g. located at the border between IPv4 Internet and IPv6 SANDRA ground networks), mapping IPv4 CoA to an IPv6 CoA for NAPT-PT.

5.2 Header compression

The network architecture which is driven by the security, mobility, and quality of service requirements, imposes the addition of a significant amount of header to the original message payload. Fig. 4 shows the structure of packets sent to the air interface at the egress of the IR. A quick observation shows that the amount of overhead justifies the need for implementing a header compression mechanism in particular for low bandwidth links.
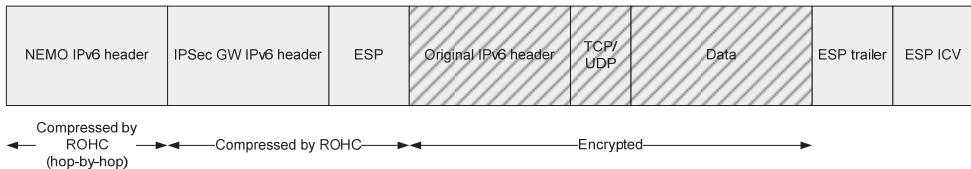

Fig. 4. Packet structure going out from airborne Integrated Router.

Plenty of header compression schemes currently exist and are standardised in various IETF RFCs. For the SANDRA testbed, Robust Header Compression (RoHC) (Pelletier & Sandlund, 2008) is chosen because of its flexibility and the well known good performance also over long-delay, error prone links, such as satellite links. Further, RoHC is the header compression scheme considered in the ICAO ATN IPS SARPS (ICAO Standards and Recommended Practices) for optional implementation in IP nodes (ICAO, 2010).

RoHC could generally be deployed between several instances[4]

- between the IR and the HA to reduce overhead of the IPv4 and IPv6 packets tunnelled by NEMO. Note: compression gains, respectively, implementation complexity is

---

[4] However, nesting RoHC is not recommended and it needs to be investigated how this issue can be addressed in the SANDRA testbed.

significantly affected by IPsec deployment, in particular if encryption is used (cf. next bullet)

- if IPsec is deployed the relatively new RoHCoIPsec related RFCs could be implemented (Ertekin et al., 2010a, 2010b, 2010c) and deployed in the respective IPsec gateways,
- between the IR and the AR, e.g. within a point-to-point protocol (PPP) context, if supported by the AR.

RoHC requires access to IP and TCP/UDP header for compressing packets. When these headers are encrypted (e.g. when using IPsec ESP) the compression is not possible. RoHC has defined a profile for ESP, but it only enables the compression of IP and the 8-Byte ESP header. Moreover, the compression works only in a hop-by-hop basis. The encrypted inner header remains untouched, as indicated in Fig. 4.

A framework for integrating RoHC over IPsec (RoHCoIPsec) is defined in (Ertekin et al., 2010a). It defines the modification required in IPsec protocol stack to allow the compression of IPsec encrypted packets.

Further investigations will be performed in the frame of the testbed development to assess in detail whether the expected compression gain obtained by implementing RoHCoIPSec justifies the implementation effort. Also alternatives to RoHCoIPsec will be investigated, such as Multilayer-IPsec (Zhang, 2004).

Due to regulatory constraints, ANSPs (Air Navigation Service Providers) do not allow data traffic in their network to be encrypted. This puts ATS, AOC traffic in a particular situation, where IPsec with AH or ESP Null encryption could be deployed for ensuring domain separation, authenticity, and integrity of the packets, but not their confidentiality. As the inner IPv6 and transport layer header are thus unencrypted, a dedicated RoHC profile could be defined to compress all the headers except the outermost one, as it is needed for routing.

5.3 ATS and AOC data traffic

Performance assessment in the testbed, not only for the assessment of RoHCoIPsec and RoHC for AH or ESP Null encryption (cf. previous section), requires realistic ATS, AOC, AAC, and APC data traffic.

For ATS and AOC, the services defined in the COCR (Communications Operating Concept and Requirements for the Future Radio System) will be used (EUROCONTROL/FAA, 2007). A challenge is to obtain a communication traffic profile per aircraft (i.e. packet length, and frequency of occurrence), imposed by the COCR messages to the network. A realistic estimate needs to take into account the COCR service instance, which states, for example, that a message exchange shall take place once per sector, once per domain in airport, etc.

Our COCR traffic estimate is obtained from the methodology and tool developed in the framework of the EC project ANASTASIA (Airborne New and Advanced Satellite Techniques & Technologies in A System Integrated Approach) and ESA Iris programmes. The methodology uses information from a global flight route simulation. The obtained information, e.g. average flight duration, average flight duration in a domain, etc., is used to approximate and normalise the COCR service instances, thus simplifying the complex multi-service COCR traffic.

The methodology has been presented to the (satellite) ATM community several times, including EUROCONTROL and NexSAT meetings in particular, and received comments have been continuously used to improve methodology and tool. The average traffic intensity of COCR messages (in bit-per-second per aircraft) will be used, e.g. in the context of deriving RoHCoIPSec compression gain.

5.4 Consideration on transport layer and IP fragmentation

Some networking design aspects are also considered in the design of the SANDRA testbed. Transport layer design is currently being investigated within the SANDRA project. Basic TCP is not efficient to use due to some issues specific to the aeronautical communications:

- *TCP slow start mechanism*: TCP slow-start phase takes several RTTs to achieve the optimum throughput. As most services require only few exchanges of small messages (some tens of kB), the available link capacity will be used inefficiently.
- *TCP congestion control algorithm*: TCP interpret segment losses as an indication of link congestion, which is usually the case in the terrestrial Internet, and immediately reduces its congestion window to reduce the network load. In wireless links bit errors may also cause segment loss, in which case reducing window size (and hence throughput) is not necessarily the most appropriate strategy.
- *Different links with different characteristics*: The variety of the link technologies also poses a challenge, because optimisation of TCP parameters with respect to one link may not be applicable to other links.

Several alternative solutions have been investigated including:

- Solutions to TCP slow start issue: increasing TCP Initial Window (through TCP parameters configuration), and TCP start-up optimisation mechanisms such as Jump Start, Quick Start, and Swift Start.
- Congestion control alternatives: TCP CUBIC, Fast-TCP, Compound TCP, and Explicit Congestion Notification (ECN).
- ECN extensions: eXplicit Congestion Protocol (XCP) and Rate Control Protocol (RCP).

Based on the earlier investigation within NEWSKY, two separate approaches are considered for safety and non-safety domains: an enhanced TCP solution is recommended for safety domain, and a PEP-based solution for the non-safety domain. XCP is currently considered to be an attractive TCP enhancement approach. The SANDRA testbed takes into account the transport layer issue and will implement the most appropriate solution accordingly.

Another issue is related to Maximum Transmission Unit (MTU) and IP fragmentation. Although generally not considered as an issue in IPv6, IP fragmentation deserves also some consideration. In a tunnelled connection between two routers, the tunnel end points become the communication end points. RFC 2473 specifies that the routers are allowed to fragment the packets if its size is larger than the tunnel MTU (i.e. the link/path MTU minus the tunnel overhead(s)), and if the tunnel MTU is smaller than the minimum IPv6 MTU requirement (1280 Byte).

One such scenario is when the IMR adds the mobility header, causing the packet size to be larger than the link MTU, e.g. satellite link, and fragments the tunnelled packet before sending it to the link. The fragmented packets can then be only reassembled at the HA. The issue of in-network reassembly then applies, e.g. the HA needs to allocate sufficient buffer to anticipate lost or out-of-order fragments, and if a fragment is lost the whole packet is considered as lost. These issues can cause a significant routing slowdown at the HA, and in turn adding up to the end-to-end latency. The SANDRA testbed will take this into account and implement the necessary solutions whenever applicable, e.g. adjusting the MTU values to allow for maximum possible packet size, or by using Path MTU Discovery (PMTUD) protocol.

6. Conclusions

This chapter has presented a detailed overview over the SANDRA testbed, including a comparison with its precursor, the NEWSKY testbed.

It was stated that the main difference between the two testbeds is that the SANDRA testbed will represent a *proof-of-principle prototype* of a future aeronautical communications system, whereas the NEWSKY testbed resembled a *proof-of-concept*.

This has quite far reaching consequences in terms of complexity of the testbed design and implementation efforts and also for verification&validation and trials, which need to be performed in a relevant environment, comprising integration on an aircraft and performing in-flight trials.

In contrast to the NEWSKY proof-of-concept, the SANDRA prototype needs a high degree of automatism for QoS provision, dynamic link selection, and resource allocation, etc. These functionalities are jointly implemented in SANDRA by the Integrated Router and the Integrated Modular Radio, which both were not present in NEWSKY.

Detailed design and implementation of the SANDRA testbed components including the SANDRA prototypes will continue until 2nd quarter 2012. Subsequently, all testbed components will be integrated, followed by extensive laboratory trials (until 1st quarter 2013) to verify&validate the SANDRA testbed and to investigate the performance of the deployed prototypes and protocols.

After successful integration of the testbed, flight trials on an Airbus A320 will be performed in the 2nd quarter 2013 to assess performance in a real-world scenario.

Of course, we intend to publish the final testbed design and trials results in future publications.

7. Acknowledgement

8. References

Ayaz, S.; Bauer, Chr.; Arnal, F. (2009). Minimizing End-to-End Delay in Global HAHA Networks Considering Aeronautical Scenarios, *Proceedings of the 7th ACM international symposium on Mobility management and wireless access (MobiWAC '09)*, Tenerife, Canary Islands, Spain, Oct. 26-27, 2009.

Devarapalli, V.; Wakikawa, R.; Petrescu, A. & Thubert, P. (2005). Network Mobility (NEMO) Basic Support Protocol (RFC 3963), Available from: datatracker.ietf.org/doc/rfc3963

Ertekin, E.; Jasani, R.; Christou, C. & Bormann, C. (2010a). Integration of Robust Header Compression over IPsec Security Associations (RFC 5856), Available from: datatracker.ietf.org/doc/rfc5856

Ertekin, E.; Jasani, R.; Christou, C.; Kivinen, T. & Bormann, C. (2010b). IKEv2 Extensions to Support Robust Header Compression over IPsec (RFC 5857), Available from: datatracker.ietf.org/doc/rfc5857

Ertekin, R.; Christou, C. & Bormann, C. (2010c). IPsec Extensions to Support Robust Header Compression over IPsec (RFC 5858), Available from: datatracker.ietf.org/doc/rfc5858

EUROCONTROL/FAA (May 2007). *Communications operating concept and requirements for the future radio system* (COCR Version 2.0)

Fazli, E. H.; Via, A.; Duflot, S. & Werner, M. (2009). Demonstration of IPv6 Network Mobility in Aeronautical Communications Network, *Proceedings of International Conference on Communications 2009 (ICC 2009)*, Dresden, Germany, Jun. 14-18, 2009

ICAO (2010). *ICAO Doc 9896, Manual on the Aeronautical Telecommunication Network (ATN) using Internet Protocol Suite (IPS) Standards and Protocols* (1st Edition), ICAO, ISBN 978-92-9231-439-2, Montréal, Quebec, Canada

Johnson, D.; Perkins, C. & Arkko, J. (2004). Mobility Support in IPv6 (RFC 3775), Available from: datatracker.ietf.org/doc/rfc3775

Nautilus project web page. (2009). 03.5.2011, Available from: www.nautilus6.org

Pelletier, G. & Sandlund, K. (2008). RObust Header Compression Version 2 (RoHCv2): Profiles for RTP, UDP, IP, ESP and UDP-Lite (RFC 5225), Available from: datatracker.ietf.org/doc/rfc5225

Perkins, C.; Johnson, D. & Arkko, J. (2011). Mobility Support in IPv6 (IETF Internet Draft), Available from: datatracker.ietf.org/doc/draft-ietf-mext-rfc3775bis

SANDRA project web page (2011). 03.5.2011, Available from: www.sandra.aero

Soliman, H. (2009). Mobile IPv6 Support for Dual Stack Hosts and Routers (RFC 5555), Available from: datatracker.ietf.org/doc/rfc5555

Tsirtsis, G.; Soliman, H.; Montavont, N.; Giaretta, G. & Kuladinithi, K. (2011). Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support (RFC 6089), Available from: datatracker.ietf.org/doc/rfc6089

Via, A.; Fazli, E. H.; Duflot, S.; Riera, N. & Werner, M. (2009). IP Overhead Comparison in a Test-bed for Air Traffic Management Services, *Proceedings of the Vehicular Technology Conference 2009 (VTC 2009)*, Barcelona, Spain, Apr. 26-29, 2009

Wakikawa, R.; Devarapalli, V.; Tsirtsis, G.; Ernst, T. & Nagami, K. (2009). Multiple Care-of Addresses Registration (RFC 5648), Available from: datatracker.ietf.org/doc/rfc5648

Zhang, Y. (2004). A Multilayer IP Security Protocol for TCP Performance Enhancement in Wireless Networks. *IEEE Journal on Selected Areas in Communications*, Vol.22, No.4, (May 2004), pp. 767-776, ISSN 0733-8716

Part 3

Challenges for the Satellite Component

The Role of Satellite Systems in Future Aeronautical Communications

Nicolas Van Wambeke and Mathieu Gineste
Business Segment Navigation and Integrated Communications,
Thales Alenia Space
France

1. Introduction

Recent evolutions in the context of aeronautical communications have changed the landscape and the role of the different systems that allow aircrafts to maintain a link with the ground while in flight. The increase in capacity needed to support the growth of worldwide air traffic and the need for increased communication safety are driving a transition from voice-centric procedures aided by slow data link connections to data-centric control applications executed on higher capacity communication systems. These future data links have to fulfil very stringent performance requirements. Indeed, the nature of the information they carry which is bound to become the first mean of air traffic control make their availability critical to the safety of air transportation in the future.

Satellite communication systems have many differentiating arguments when compared to terrestrial solutions. Indeed, while the deployment costs of terrestrial systems can be sustainable in high-density areas, their use in low-density remote areas is much less interesting. In high-density areas, satellite could also be useful either as a primary mean of communication or as a secondary one in order to improve the overall communication system's availability. A satellite system, by nature, is able to cover large regions of the earth and can thus provide a cost effective solution to the coverage of both high and low density areas such as oceanic regions where reliable terrestrial coverage is nonexistent.

In this paper, the interest of a satellite solution to be used as a data link for future aeronautical communications is studied. After presenting an overview of the existing satellite systems for aeronautical communication in operation today, the discussion focuses on the interest and strength of the forthcoming satellite link under definition in the frame of the ESA Iris Programme and its integration in the communication concept defined by the SANDRA EC FP7 project.

2. Designing satellite systems for aeronautical communications

This section presents the characteristics of satellite systems supporting aeronautical communications. A first part of the study presents the general architecture of such a system as well as its integration in the communication infrastructure as an access network to the aeronautical telecommunications network. In a second part, the constraints imposed by the regulatory domain on spectrum usage are presented. Finally, the most characteristic systems in place today are presented.

2.1 General architecture of a satellite aeronautical communication system
In this section, the general architecture of a satellite system for providing aeronautical communication services is presented. After presenting the role of the different segments, the integration of such systems in the aeronautical telecommunications network (ATN) specified in (International Civil Aviation Organization [ICAO], 2007) is presented.

2.1.1 Ground, space and user segments
A typical satellite communication system is divided into three different segments. These are respectively the ground, space and user segments.

The ground segment is responsible for interfacing the satellite communication system with the rest of the communication network infrastructure to which the satellite system constitutes an access network. Indeed, networking infrastructures are structured with a core network to which several access networks interconnect in order to allow end users to connect. In the ground segment of a satellite communication system, the information stream that arrives through the ground infrastructure is adapted in order to be sent out on the air interface of the satellite network gateway, which in aeronautical communication systems is known as a ground earth station (GES).

The space segment is composed of the satellite itself, the role of the space segment is to either serve as a transparent reflector for the signals sent from the ground or to receive process and re-generate a signal towards the ground in which case the satellite is called regenerative. A regenerative satellite can be used in the case where the equipment on the ground and user segments doesn't use the same modulation and coding rate for example. Another example of regenerative satellites are those used in constellations such as Iridium for which the signal is decoded in the space segment in order for it to be routed towards the appropriate satellite towards its destination.
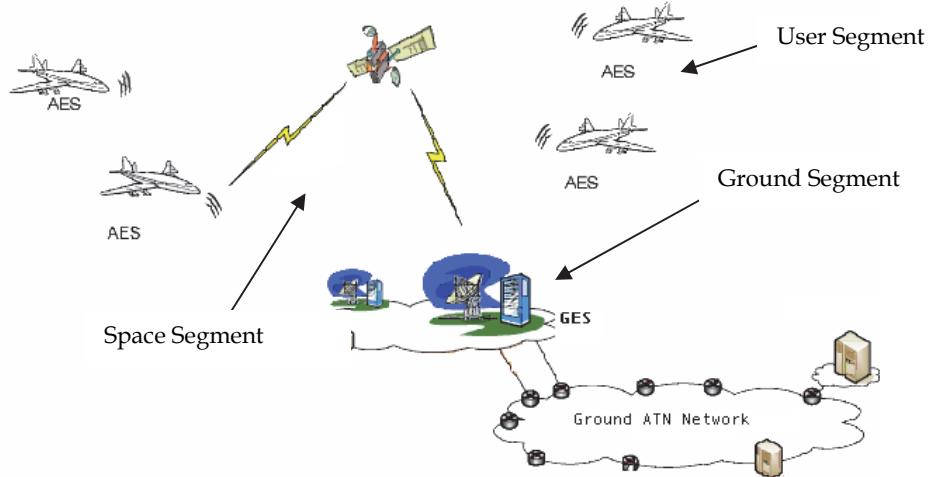

Fig. 1. Satellite system for aeronautical communications architecture as considered in (ICAO, 2010).

The user segment as its name indicates is where the users of the satellite communication system are located. In the case of an aeronautical communication satellite system, the user

segment is known as the aeronautical earth station (AES). The role of the user segment is to provide an interconnection mechanism between the on-board networks and systems and the satellite access network. In a way that is similar to the ground segment, the user segment provides the interface between the streams of data that are under the control of the satellite system (implementing a specific communication standard) and the outside world.

2.1.2 Satellite system integration to the aeronautical telecommunications network

A huge growth of aeronautical traffic is foreseen in the next few years. Predictions state that by the year 2020 to 2030, a new paradigm in aeronautical communications has to be envisaged to cope with this increase by defining new Air Traffic Management (ATM) concepts.

EUROCONTROL and the Federal Aviation Administration (FAA) have initiated a joint study reported in (EUROCONTROL/FAA, n.d.) to identify potential future communications technologies to meet safety and regularity of flight communications requirements, i.e. those supporting Air Traffic Services (ATS) and safety related Aeronautical Operational Control (AOC) communications.

The objective is to replace progressively voice communication for air traffic management by data communications services for safety reasons and because it supports increased automation in the aircraft and on the ground. Potential resource savings should also be possible when replacing voice by data communications. These data communications should then become the primary means for safety air-ground communication.

These data link oriented communication services will be supported by new communication infrastructures. On the ground, a core aeronautical telecommunications network will be used to interconnect the various ATC and AOC centres together. Furthermore, several technology specific access networks (i.e. Satellite, LDACS, AeroMACS, …) will allow the aircrafts to form part of the ATN.

Fig. 2 provides a view of an end-to-end communication handled by a satellite aeronautical communication system. This view highlights both physical architecture, and logical architecture mapped on a network layers definition.

The satellite system includes the AES (aircraft on board terminal), and the GES (Earth terminal); this system is depicted in red line on Fig. 10.
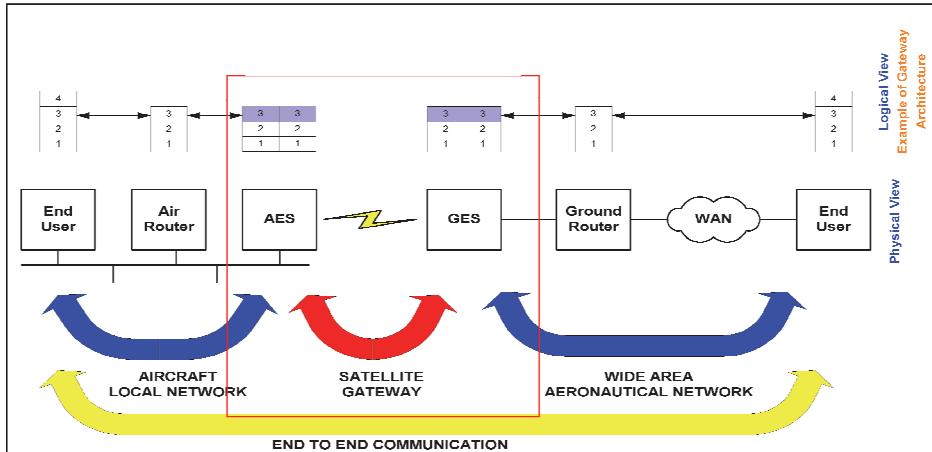

Fig. 2. SATCOM system integration in the ATN network.

The satellite is not represented in this figure, because the baseline architecture considers that the satellite payload is "transparent". This means that, functionally, no processing exists on board the satellite, neither on the data nor on the frames; such a payload only handles a frequency conversion function.

It is interesting to note that the network layer, namely layer 3 is greyed out on the above figure. Indeed, a satellite system can either operate at layer 3 and thus appear as being an active element of the network (an IP router) or it can operate completely at layer 2 in which case it constitutes a transparent bridge equipment between several segments of the same IP network.

In future satellite systems for aeronautical communications, it is foreseen that their operation is performed at layer 3 for reasons which are linked to mobility requirements among others which are further detailed in section 3.1 hereafter.

2.2 Regulatory constraints on spectrum usage

Aeronautical communication systems used for the transport of ATC/AOC are considered as safety critical in their frequency allocation by the ITU while systems used for APC communications are not.
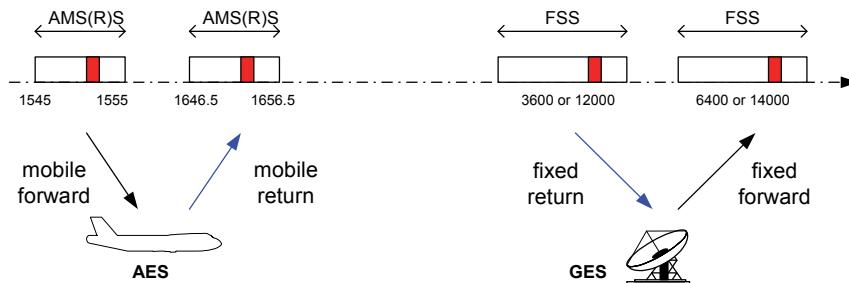

Fig. 3. Typical frequency bands used for safety aeronautical communications via satellite. These band allocations are managed by the ITU.

The principle of transmission in the safety satellite system is shown in Fig. 3:
- The mobile link, between the satellite and the aircraft, is built on a safety satellite spectrum allocation, based on AMS(R)S standard;
- The satellite is in charge of signals frequency conversion, simultaneously from C or Ku band to L band for the forward link, and from L band to C or Ku band for the return link;
- The fixed link, between the ground and the satellite, is built on a fixed satellite spectrum allocation, based on FSS standard.

In following sections, the regulatory situation in the L band for safety and non-safety services is exposed, as well as for the Ku band.

2.2.1 L band situation

The L band is defined as the Mobile Satellite Service allocation in the frequency ranges 1525-1559 MHz and 1626.5-1660.5 MHz.

Although the whole band is generically for Mobile Satellite Service (MSS) use, in certain portions of the band, safety related services are afforded a specific status in the ITU radio regulations, as shown on Fig. 4.

In the sub-band 1646.5-1656.5 MHz and 1545-1555 MHz, the communications in the AMS(R)S are afforded priority over other types of communications, through the footnote 5.357A of the Radio Regulations (International Telecommunications Union [ITU], 2008).
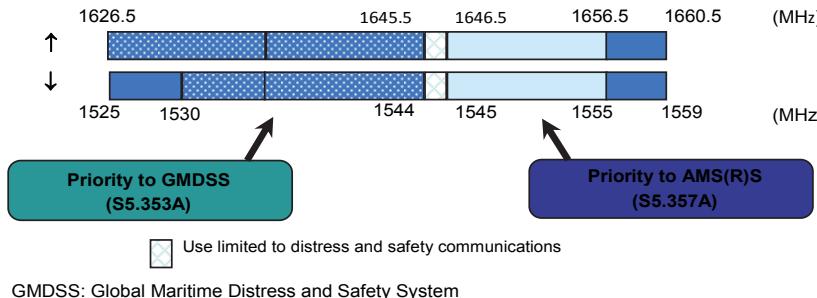

Fig. 4. AMS(R)S L band allocation for SATCOM.

The concerned communications are those falling under categories 1 to 6 of Article 44 of the Radio Regulations, as listed below:

1. Distress calls, distress messages and distress traffic.
2. Communications preceded by the urgency signal.
3. Communications relating to radio direction finding.
4. Flight safety messages.
5. Meteorological messages.
6. Flight regularity messages.
7. Messages relating to the application of the United Nations Charter.
8. Government messages for which priority has been expressly requested.
9. Service communications relating to the working of the telecommunication service or   to communications previously exchanged.
10. Other aeronautical communications.

In the specific context of the L band, given the technical nature of satellite systems involved, it has been felt more efficient to have multilateral meetings among the concerned parties instead of solely relying on Article 9 of the Radio Regulation (ITU, 2008). In effect, the terminals in the L band have poor directivity which impacts any satellite network operating in visibility of that terminal, which leads to segmentation of the spectrum among systems.
Given the high demand for spectrum in the L band, it is difficult for a new entrant to have spectrum granted, even if this is for safety services. For non-safety services it is seen as impossible to have significant spectrum allocated for a new entrant.

2.2.2 Ku band situation

In this section, the discussion is limited to the downlink portion of the Ku band, i.e. the portion dedicated to the reception by the aircraft.
The downlink portion of the Ku band is divided in allocations to various services:
-   FSS planned band: These bands are regulated by Appendix 30B of the Radio Regulations. In these bands, every country member of the ITU has access to a reserved orbital position for national coverage. There are few operational systems in this band

- FSS unplanned bands: these bands host most of the Ku band satellite systems, because of the flexibility of its regulation. In many areas, this band is fully occupied by operational systems.
- BSS planned bands: these bands are regulated by Appendix 30 of the Radio Regulations (RR, 2008). In these bands, every country member of the ITU has access to a reserved orbital position and determined number of TV channels for national coverage. There is a significant number of operational systems in this band
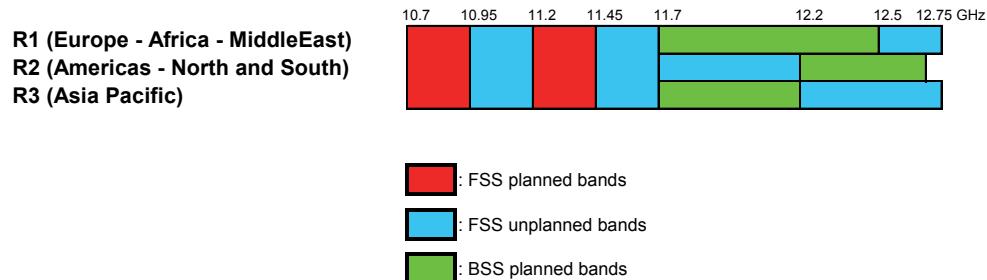

Fig. 5. Ku band allocation for SATCOM downlink.

The Ku band downlink allocations are depicted in the diagram presented on Fig. 5. Most of operational Ku band satellites operate in the Ku FSS unplanned band (in blue). Therefore it is likely that the capacity for APC services could be found in this portion of the spectrum, although the others are not excluded.

It should be noted that it would be more correct from a regulatory point-of-view to use Mobile Satellite Service allocations for the service since an aircraft can be considered as a mobile terminal. However, it is possible to use FSS allocations, as long as the services to not entail regulatory requirements higher than those of a classical FSS use.

Regarding inter-service sharing, the band 10.7-11.7 GHz is shared with terrestrial Fixed and Mobile services worldwide, with a majority of fixed use. In Region 3, the band 12.2-12.75 GHz is also shared with terrestrial services, as well as in Middle East and Africa for 12.5-12.75 GHz.

In order to enable sharing, there are power flux density limits applying to satellite transmissions in these bands (In these shared bands, the satellite systems for aeronautical communications receiver may experience bursty interference events due to fixed links interference).

Intra-service sharing concerns the potential interference among satellites systems sharing the same band. In order to maximize the use of the orbit, the ITU has developed recommendations on off-axis gain of earth stations (ITU, 2003), (ITU, 2010), (ITU, 2010a)

In satellite systems for aeronautical communications, the Ku band would be used in Receive only mode, therefore without risk of interference from the earth stations towards other systems.

2.3 Frequency allocation and spot beams definition

Considering the restrictions presented above, the different aspects of system dimensioning that need to be taken into account when designing an aeronautical communication system

are presented. Indeed, the use of several carriers, the frequency at which these carriers operate but also the geographical extend of these frequency domains on the ground (known as spot beams) are elements that need to be specifically adapted to the aeronautical context.
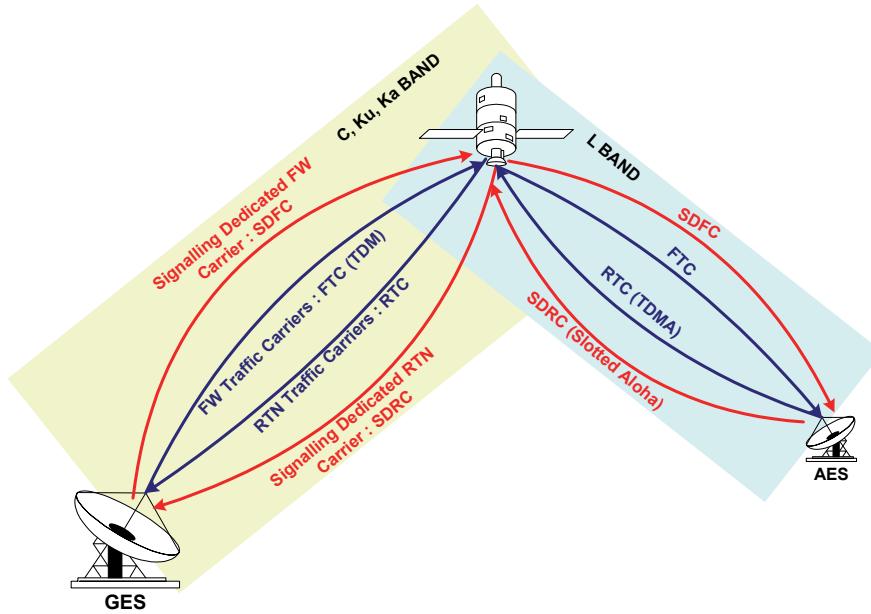

Fig. 6. Typical channel allocation.

A typical channel allocation between the GES and the AES for aeronautical communication systems via satellite is presented on Fig. 6, as it can be seen, the feeder link channels (between the GES and the satellite) are operated in C, Ku or Ka bands while the user link (between the satellite and the AES) are operated in the L frequency band.

Typical satellite architecture for aeronautical communications should include the following frequency carriers:

- Global Beam (meaning a common carrier on forward and return link for the whole system) shall be used for initial signaling meaning for initial system information, log-on and initial synchronization.
- The carriers for the global beam are named SDFC (Signaling Dedicated Forward Carrier) for the Forward Link and SDRC (Signaling Dedicated Forward Carrier) for the Return Link on the figure herein.
- On the opposite, within a spot beam one or more carriers can be used and are named FTC (Forward Traffic Carrier) and RTC (Return Traffic Carrier).

This architecture is summarized on Fig. 6, while these carrier names are not standard, they are logically implemented by most systems used for aeronautical communications.

Spot beams are required in the system to use the capacity more efficiently. A spot beams hypothesis is important in particular when the frequency allocation plan is realized. Spot beams also have the advantage of enabling the system to have a pattern for re-using frequencies. This permits to have more capacity on the global system given that one frequency is used several times in the global coverage.
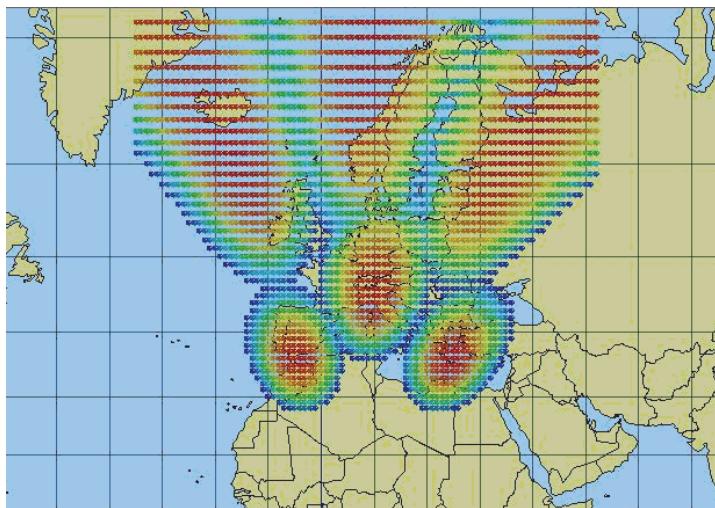
Fig. 7. Illustration of the interest of using spot beams in satellite system. In this example, when the overall available spectrum is limited, the gain provided by frequency reuse between non-adjacent spots can be visually verified.

2.4 Existing satellite systems in operation for aeronautical communications
2.4.1 Inmarsat and MTSAT

The Inmarsat service was initially targeted to providing a maritime communication service to the community for safety of life related issues. However, Inmarsat soon began to provide service to other communities such as aircraft and mobile users.

The space segment of the Inmarsat system is a constellation composed of several geostationary satellites (the number of satellites depend on the service as not all of them support all the services) that cover the earth with the exception of the poles. Aeronautical services supported by the system are currently ATS and AOC services. These can either be used through the legacy ClassicAero service or the recently introduced SwiftBroadband service (based on the BGAN technology adapted to the aeronautical context). The Inmarsat satellites use three different types of spot beams, one global spot beam for initial signalling and specific services, a set of regional spot beams (since the 3rd generation satellites) and very small spot beams (radius in the order of hundreds of kilometres) used for the BGAN service and allowing for smaller antennas to be used on the handheld terminals.

In terms of frequencies, the Inmarsat system operates the feeder link in Ku bands and the user link in AMS(R)S reserved portions of the L band.

The Classic Aero service is mainly used for establishing circuit oriented connections for low and medium quality voice and fax. In addition to these services, packet data services such as ACARS and ADS can also be used.

The SwiftBroadband service offers much higher data rates than Classic Aero and takes advantage of the small spot beams of the fourth generation satellites to provide users with these data rates. The SwiftBroadband service is based on the use of the IP protocol at network layer and is mainly used to provide passengers with Internet access.

In addition to the Inmarsat satellite constellation, the Classic Aero protocol is also used by the MTSAT system operated for the Japanese Civil Aviation Bureau (JCAB). The MTSAT

system as described in (Oikawa & Kato, 2006) offers ATS and AOC services to airlines in the Asia/Pacific area and provides increased availability by using two specifically located geostationary satellites (MTSAT-1R and MTSAT-2).

2.4.2 Iridium
In addition to the Inmarsat and MTSAT satellite systems presented above, the Iridium low earth orbit constellation of telecommunication satellites also provides aeronautical communication services.

The Iridium constellation is comprised of 66 active satellites that provide complete coverage, including the earth poles. The feeder and inter-satellite links are operated in Ka frequency band while the user link is operated in the L band.

Services offered by the Iridium constellation are based on the GSM standard and include both voice and data oriented communications. In addition to these services, one-way paging services are also possible.

The Iridium constellation and services has recently been undergoing the authorization process required to be used for AMS(R)S services. However, initially, the system will be used to provide voice-oriented communication between controllers and pilots for the needs of ATS services.

3. The future satellite link: challenges

3.1 Overview of the future aeronautical communication infrastructures
Currently, in continental areas, ATM mobile communications use a narrowband VHF (Very High Frequency) voice system combined with a VHF digital data link, e.g. VDL (VHF Digital Link) Mode 2 (Fig. 8 represents a VDL network architecture). The VHF network is composed of terrestrial antennas connected with gateway routers to a backbone network in which the services are located. Although VHF is a very mature and reliable technology, it presents some disadvantages. It requires several remote ground stations to achieve the coverage that implies high operating cost due to links between ATC centres and remote radio stations. The coverage is limited to line of sight so the number of required station increases in non-flat areas.
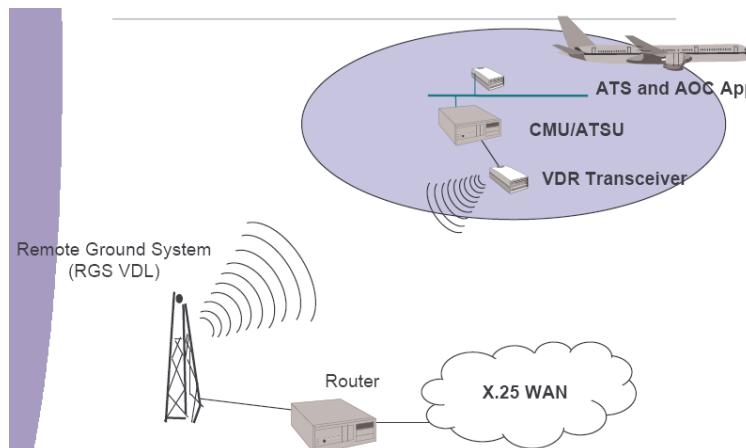

Fig. 8. VDL architecture.

In remote areas and over oceans, HF (High Frequency) and SATCOM (SATellite COMmunications) voice and data link systems are used. HF network has the same architecture as the VHF one but it is not limited to line of sight propagation, it can also be used with ground wave propagation and sky wave propagation (through reflexions on various atmosphere layers). The main drawback of HF communications is its poor link overall quality due to fading. HF tends to be replaced by satellites links in oceanic areas because of the higher quality of satellites communications. But the currently implemented satellites links are not efficient enough to be economically viable on a large-scale deployment. Fig. 9 depicts the general architecture of ATM network.
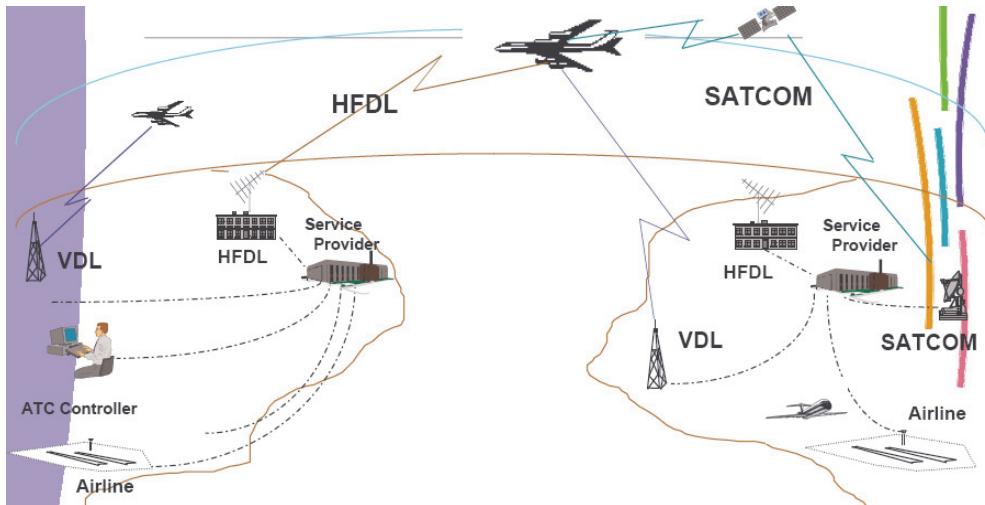

Fig. 9. ATM network global architecture.

The evolution from voice to data links for ATC is motivated by safety reasons and by saturation of voice links in dense areas. Indeed, data transmission allows using less bandwidth and safer communications. With the considerable increase of the air traffic last few years and the expected increase in the next coming years the current ATM network will not be able to handle all the traffic with the requirements associated to ATC services. In dense area managed airspace the objective is thus to increase ATM capacity while having even higher level of safety and getting rid of aeronautical routes. In low density managed airspace the objective is to have a higher communication quality and more flexibility in trajectories of aircrafts.

Future aeronautical communication architectures, which are currently being defined by initiatives and programmes in both Europe (through the SESAR programme) and the USA (through the NextGen programme), will allow for onboard end systems to communicate with other end systems located on the ground through potentially more than one radio link at a given time. From a topology point of view, this functionality can be illustrated as shown on Fig. 10.

On the airborne side of the network, several functional entities are represented, from passenger end systems which will mainly use the network architecture in order to access the Internet and specific passenger services to ATS/AOC applications which will communicate

with ATS/AOC service providers on ground through the use of potentially multiple access networks technologies including the satellite link.

On the ground side of the network, the counterpart to several of the airborne side functional entities are presented. Indeed, in order to provide its service, the network architecture relies on functionalities provided on the ground by Mobility Information Services as well as Security Services. Finally, the figure also presents the ATS/AOC service providers, which are the onboard systems and applications counterparts.
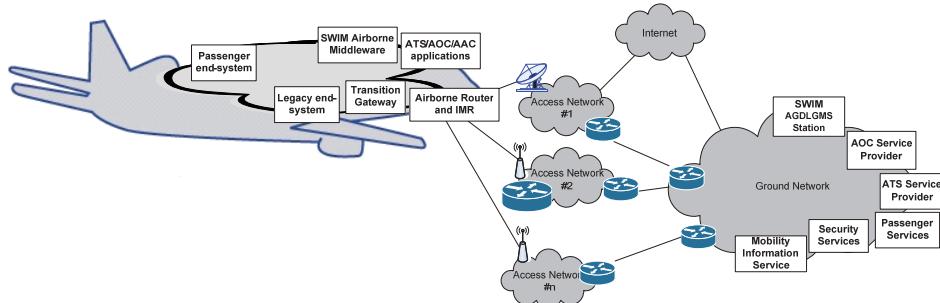

Fig. 10. Example of satellite link integration in the Future Aeronautical Communications infrastructure.

In order to maintain a global connectivity between airborne and ground networks, the network layer shall be implemented using the IPS protocol suite (ICAO, n.d.) which is strongly inspired by the IETF defined IPv6 protocol (Deering & Hinden, 1998). Furthermore, the onboard equipments are located on an IPv6 network that can be considered as being a mobile network according to the definition provided by the Mobile IPv6 standards (Johnson et al.). Indeed, throughout a flight, the airborne router (and the airborne network behind it) might from one access network to another (i.e. a switch from terrestrial LDACS to a satellite link during a transatlantic flight). In addition to being mobile, future architectures foresee that the airborne router establishes connections to multiple access network technologies at the same time. In this case, the network mobility extensions to Mobile IPv6 (Devarapalli et al., 2005) have to be complemented by specific strategies (Ng et al., 2007) to handle these multiple links.

Fig. 11 illustrates a possible instantiation of the previously described situation. In this scenario, the mobility tunnels are established between the airborne router and the home agent through each of the available data links. In this context, the loss of a given data link connection has the effect of removing one of the tunnels between the aircraft and its home agent. The advantage of this solution lies in the fact that no routing updates are required when new connections are established or existing connections are lost.

3.2 Role of the satellite systems

The current ATM network is very heterogeneous and the past evolutions of this network have been done without considering the need for global interoperability and a constant quality of services. Several networks and protocols stacks are used for different services. The next evolution will be to move toward a single data transport network supporting all the services and IP technology is the natural evolution to this objective.
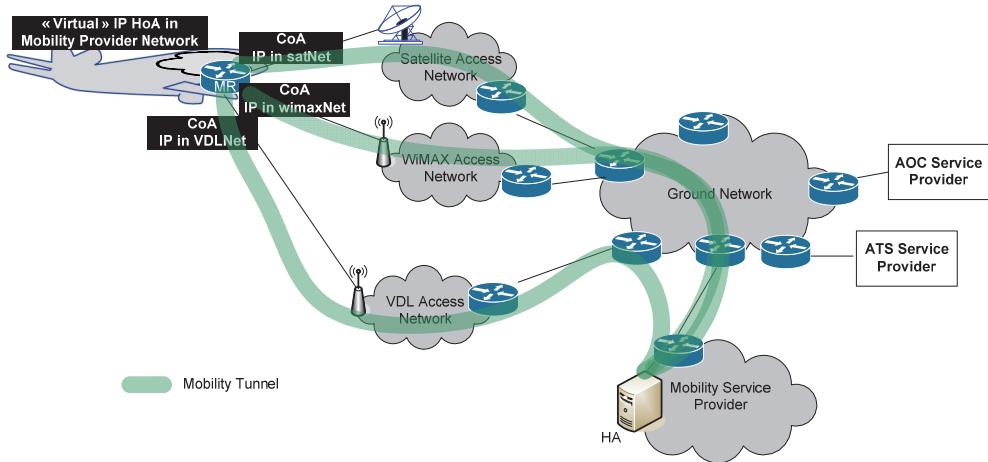
Fig. 11. Possible situation where mobility tunnels are added on top of the multiple data links to access the ATN networks. This illustration makes use of a dedicated mobility service provider that at the present time is not identified as a system actor.

On Fig. 12 is presented the synoptic of the new ATM network that was proposed in the definition Phase of IRIS (European Space Agency [ESA], 2009).

In regards to the requirements concerning future ATM network that were presented in the previous section, satellite systems have substantial assets in providing access to AES. These assets are presented hereafter.

The constant quality of service on the covered area required by such systems will be possible by the mean of a global coverage that only satellites can provide. Besides the satellite can provide high safety guarantees along with a constant deployment and maintenance cost all over the coverage (conversely to terrestrial technologies) and permits to get rid of aeronautical routes.

For all these reasons, the satellite shall play a major role in the future aeronautical communication network.

Either the satellite will be used as a primary mean for future ATM communications in all areas ensuring a constant quality of service and safety communication that could potentially be backed up by terrestrial technologies (such as LDACS), if needed.

Or, the terrestrial access could be used as a primary mean of ATM communication in the dense area managed airspace while the satellite could be used as a primary communication mean in low-density area managed airspace and as a backup communication mean in dense areas.

A last option would be to use concurrently several access technologies including satellite access to provide future ATM communications. This last solution would permit to increase the overall availability of the network while reducing handover delay in case of a technology breakdown, as all available technologies would be used in parallel. This approach could also permit to maintain a constant Quality of Service for ATM communications, as the access technology the most adapted to the QoS requirements would be used.
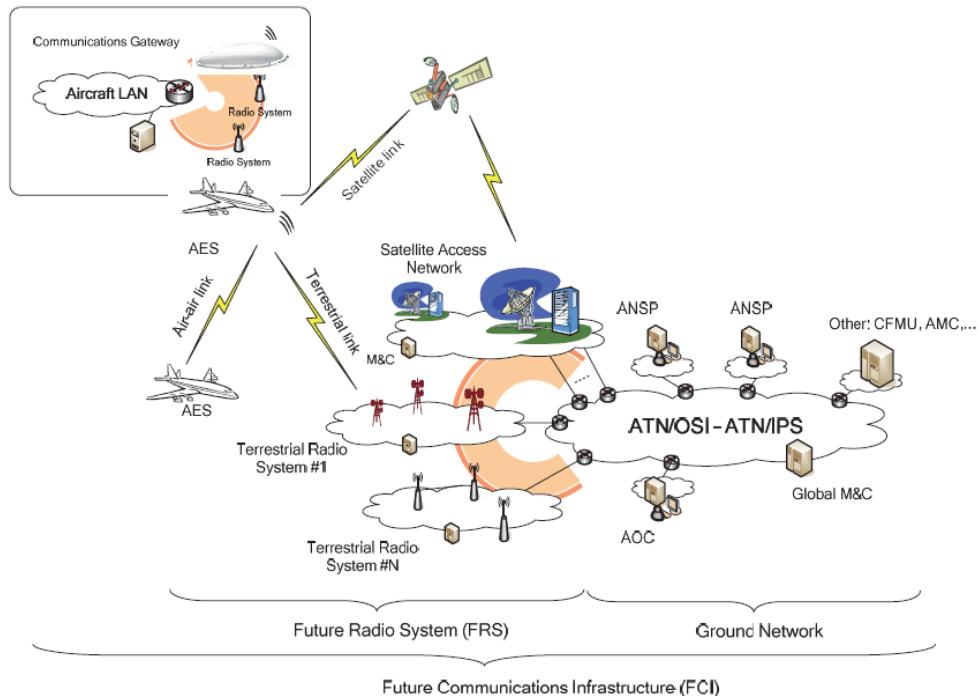
Fig. 12. Future ATM network architecture integrating the satellite access network as presented in the ICOS ESA project.

4. Conclusion

In this paper, the role of satellite systems in future aeronautical communications has been studied. Indeed, the important growth of air transportation in the upcoming years will change the way communication systems are used by pilots and crew. From a voice-centric paradigm, the convergence is likely to be a data-centric paradigm where voice is maintained only for highly critical situations for which text oriented transmissions are not adapted. In this context, the transmission delay, which is the main drawback of geostationary satellite communication systems, becomes less important than for highly interactive video/voice.

The advantages of satellite systems, on the other hand, are interesting for aeronautical communications. Indeed, the large coverage, high availability, low maintenance costs, high flexibility in resource allocation and usage as well as the ability to provide similar service in remote areas are arguments in favour of such systems.

The current programmes aiming at defining the future communication infrastructures for the ATN in both Europe and the North America are both considering the use of a satellite system as part of the access network technologies to be used. While the regulatory framework imposes some constraints on the overall capacity and system design, throughout this paper, it has been shown that a satellite component not only supports the full extend of ATC/AOC services in oceanic and remote areas but that some of the characteristics of satellite systems make them a first choice for certain services also in higher density continental regions.

5. Acknowledgment

6. References

Deering, S.; Hinden, R. (1998). *Internet Protocol, version 6 (IPv6) specifications,* RFC2460, Internet Engineering Task Force (IETF), 1998

Devarapalli, V.; Wakikawa, R.; Petrescu, A.; Thubert, P. (2005). *Network Mobility (NEMO) Basic Support Protocol*, RFC3963, Internet Engineering Task Force (IETF), January 2005

EUROCONTROL/FAA. (n.d.). *Future Communications Study Operational Concepts and Requirements Team - Communications Operating Concept and Requirements for the Future Radio System,* EUROCONTROL/FAA, Retrieved from:
http://www.eurocontrol.int/communications/gallery/content/public/documents/COCR%20V2.0.pdf

European Space Agency (ESA). (2009). *ICOS - Iris Communications System Design Study Phase-A*, ESA, Available from:
http://telecom.esa.int/telecom/www/object/index.cfm?fobjectid=28935

International Civil Aviation Organization (ICAO). (2007). *Doc. 9880-AN/466: Manual on detailed technical specifications for the Aeronautical Telecommunication Network (ATN),* ICAO

International Civil Aviation Organization (ICAO). (2010). *Doc. 9925/1*: *Manual on the Aeronautical Mobile Satellite (Route) Service*, ICAO

International Civil Aviation Organization (ICAO). (n.d.). *Doc 9896: Manual for the ATN using IPS standards and protocols,* ICAO

International Telecommunication Union (ITU). (2003). *Radiation diagrams for use as design objectives for antennas of earth stations operating with geostationary satellites*, Rec. ITU-R S.580-6, Jan 1st 2003

International Telecommunication Union (ITU). (2008). *The Radio Regulations, Edition of 2008.* ITU, 92-61-12451-8, Geneva

International Telecommunication Union (ITU). (2010). *Reference radiation pattern of earth station antennas in the fixed-satellite service for use in coordination and interference assessment in the frequency range from 2 to 31 GHz*, Rec. ITU-R S.465-6 , Jan 1st 2010

International Telecommunication Union (ITU). (2010a), *Alternative reference radiation pattern for earth station antennas used with satellites in the geostationary-satellite orbit for use in coordination and/or interference assessment in the frequency range from 2 to 31 GHz*, Rec. ITU-R S.1855, Jan 1st 2010

Johnson, D.; Perkins, C.; Arkko, J. (2004). *Mobility support in IPv6*, RFC3775, Internet Engineering Task Force (IETF), 2004

Ng, C.; Ernst, T.; Paik, E.; Bagnulo, M. (2007). *Analysis of Multihoming in Network Mobility Support,* RFC4980, Internet Engineering Task Force (IETF), October 2007

Oikawa, Y.; Kato, K. (2006). *Flight Inspection for MTSAT,* Proceedings of the 14th SIIV IFIS, Toulouse, France, June 2006

Development of a Broadband and Squint-Free K$_u$-Band Phased Array Antenna System for Airborne Satellite Communications

David Marpaung[1], Chris Roeloffzen[1], Willem Beeker[2],
Bertrand Noharet[3], Jaco Verpoorte[4] and Rens Baggen[5]
[1]University of Twente,
[2]LioniX BV,
[3]Acreo AB,
[4]National Aerospace Laboratory (NLR),
[5]IMST GmbH,
[1,2,4]The Netherlands
[3]Sweden
[5]Germany

1. Introduction

Novel avionic communication systems are required for various purposes, for example to increase the flight safety and operational integrity as well as to enhance the quality of service to passengers on board. To serve these purposes, a key technology that is essential to be developed is an antenna system that can provide broadband connectivity within aircraft cabins at an affordable price. Currently, in the European Commission (EC) 7th Framework Programme SANDRA project (SANDRA, 2011), a development of such an antenna system is being carried out. The system is an electronically-steered phased-array antenna (PAA) with a low aerodynamic profile. The reception of digital video broadcasting by satellite (DVB-S) signal which is in the frequency range of 10.7-12.75 GHz (K$_u$-band) is being considered. In order to ensure the quality of service provided to the passengers, the developed antenna should be able to receive the entire DVB-S band at once while complying with the requirements of the DVB-S system (Morello & Mignone, 2006). These requirements, as will be explained later, dictate a broadband antenna system where the beam is squint-free, i.e. no variation of beam pointing direction for all the frequencies in the desired band. Additionally, to track the satellite, the seamless tunability of the beam pointing direction of this antenna is also required. In this work, a concept of optical beamforming (Riza & Thompson, 1997) is implemented to provide a squint-free beam over the entire K$_u$-band for all the desired pointing directions. The optical beamformer itself consists of continuously tunable optical delay lines that enable seamless tunability of the beam pointing direction.

Although this particular concept of optical beamforming has been well investigated in the past (Meijerink et al., 2010; Zhuang et al., 2007, 2010), its implementation in the actual antenna system is by no means trivial. It requires extensive modelling of the antenna system, development of the system components that operate in the radio frequency (RF)

and/or the optical domains as well as the integration between these components to yield a compact and reliable system. This chapter focuses on these aspects towards the development of a full antenna system.

The remainder of this chapter is organized as follows: in Section 2 the target specifications of the PAA system is discussed. In Section 3, the concept of optical beamforming using optical ring resonators as delay elements is explained. The system modelling and performance analysis is presented in Section 4. In Section 5 the current status of the development of the components in the system is reported. In Section 6 the photonic integration scheme is discussed. The chapter closes with conclusions.

2. Antenna specifications

The intended operation of the PAA system is illustrated in Fig. 1. Antenna tiles (8x8), each consisting of 64 antenna elements (AEs), are arranged to form the total PAA system with a large number of AEs. As mentioned earlier this PAA is thus used to provide airplane passengers with live television service received from a satellite. To ensure proper signal receptions, the PAA should fulfil a set of requirements. The list of target specifications of the PAA is listed in Table 1.
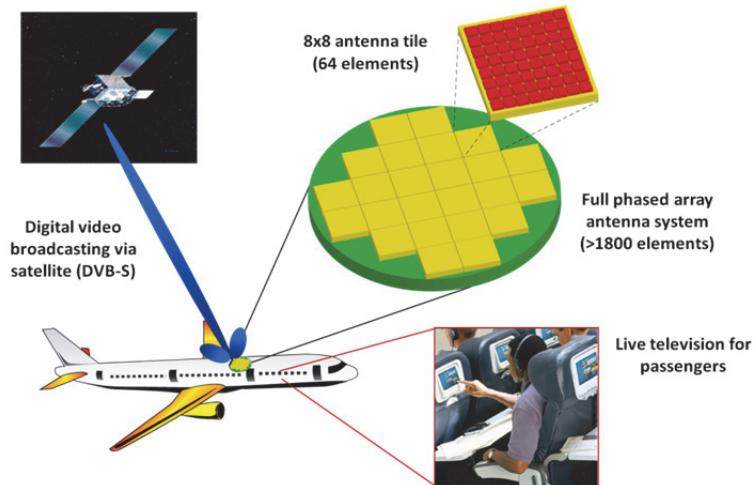

Fig. 1. Illustration of the system considered in this work. In order to provide airplane passengers with live television channels, a phased array antenna system is used for reception. The antenna consists of antenna tiles of 64 elements.

In this PAA, the received K_u-band signal is down-converted to the L-band and is subsequently processed in the beamformer. The beamformer delays and combines the signals from the AEs such that they are synchronized at the beamformer output. The time delay provided by the beamformer should be sufficient to incorporate the intended maximum scanning angle of this PAA system, which in this case is 360° in the azimuth plane and 60° in the elevation angle. Besides the maximum scanning angle, the factors that determine the required maximum time delay are the spacing between the AEs (1.18 cm in this case) and subsequently the size of the antenna. This will be explained further when the design of the beamfomer is discussed in Subsection 5.3.2.

Antenna parameter	Target value
Frequency range	10.7-12.75 GHz (Ku-band)
Frequency range after downconversion	950-3000 MHz (L-band)
Inter-element distance	1.18 cm
Steering type	Electrical, continuous
Beam steering angle	Azimuth: 360° Elevation: 60°
Beam tracking accuracy	0.1°
G/T	12.7 dB/K
Antenna size	Diameter: 60 cm Height: 10 cm
Number of elements	2048
Antenna gain	35 dBi
Polarization	Linear (H/V)

Table 1. The phased array antenna specifications targeted in this work.

The size of the antenna will also determine the antenna gain. Here, the aim is to have an antenna with a diameter of 60 cm, with a gain in the order of 35 dBi. It has been shown that this dictates at least 1800 AEs incorporated in the system (Meijerink et al., 2010). In this work we set the number of AEs to be 2048. Finally, a useful figure of merit for an antenna system is the antenna gain-to-noise temperature (G/T) which is the ratio of the antenna gain and the antenna noise temperature (in Kelvin), expressed in decibels. The G/T target value for the antenna is 12.7 dB/K. This parameter is strongly related to the antenna system architecture and will be discussed in depth in Subsection 4.1, where the system design is described.

3. Optical beamforming

The beamforming network (BFN) is a crucial part of the entire PAA system. In this work, where a broadband, squint-free and seamless beam steering is targeted, the BFN needs to provide continuously tunable time delay. In the past, many have turned to photonic delay lines to provide true time delay (TTD) (i.e. linear phase progression over the frequency) (Riza & Thompson, 1997; Meijerink et al., 2010). This is in contrast to phase shifters that provide a constant phase shift over the frequency, which will induce beam-squint for a broadband signal (Baggen et al., 2011).

Various photonic delay line structures have been proposed as the TTD element, ranging from optical fibers, fiber Bragg gratings, semiconductor optical amplifier (SOA), and integrated photonic filters. Each approach claims to offer advantage either in terms of maximum achievable delay, tunability, size and compactness, cost and/or simplicity in operation (Meijerink et al., 2010).

In the past, we have proposed a photonic BFN with optical ring resonator (ORR) filters serving as the TTD elements (Zhuang et al., 2007) which will be explained in the following subsection.

3.1 Optical ring resonator

When an optical carrier is modulated by an RF signal, propagates through an optical waveguide, and is converted to the electrical domain by an optical detector, the effective time delay to the RF signal is determined by the group delay of this optical waveguide. This group delay can be made tunable by putting an ORR parallel to the waveguide, as illustrated in Fig. 2. The group delay response is periodic and the periodicity (dubbed as the free spectral range (FSR)) is inversely proportional to the round-trip time in the ring. Each period of the group delay response is a symmetric bell-shaped function of frequency, centered at the resonance frequency of the ring (Fig. 2). This resonance frequency can be varied by tuning the round-trip phase shift, ϕ, of the ring and the maximum delay can be tuned by varying the power coupling coefficient between the optical waveguide and the ring, κ. Here, the thermo-optical tuning mechanism is used to vary the resonance frequency and the coupling coefficient of the ORR. Two chromium heaters per ORR are used for the tuning, as illustrated in Fig. 2. The principle of this ORR as delay line can be found in detail in (Zhuang et al., 2007, 2010).
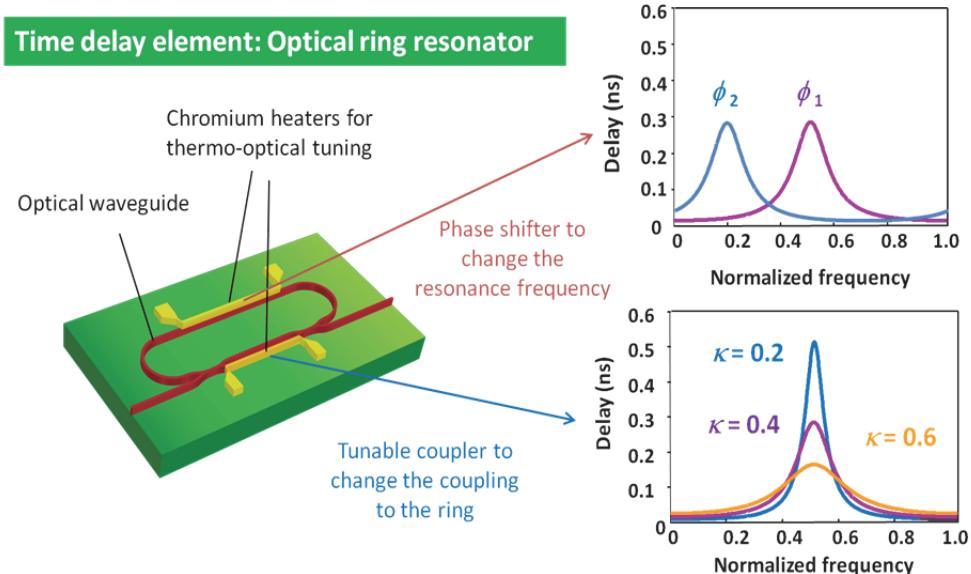

Fig. 2. Tunable optical delay based on an optical ring resonator. The group delay response of such an ORR is a bell-shaped function of the frequency, where its resonance frequency and maximum delay can be tuned using thermo-optical tuning mechanism.

The peak value of the delay is approximately inversely proportional to the width of curve since the area underneath the delay curve in one period is always constant. This imposes a trade-off between the highest delay values that can be provided while keeping the sufficient bandwidth. To overcome this, several ORRs can be cascaded, where the total group delay response is the sum of the individual ring responses. This is illustrated in Fig. 3.

In the next subsection, the principle of cascading the ORRs as tunable delay elements is used to yield the photonic BFN chip.
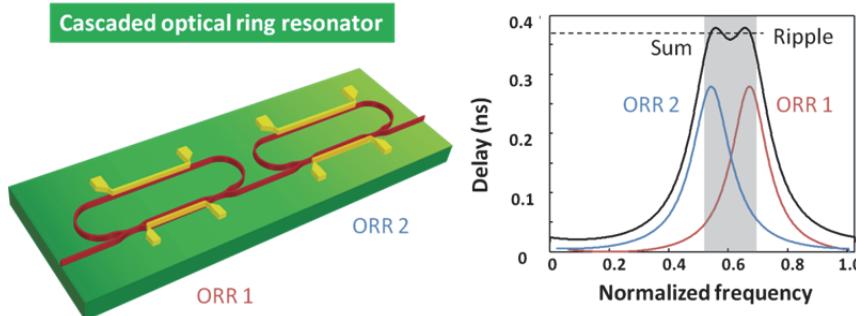
Fig. 3. Left: schematic of two serial ORRs, right: group delay response of a cascade of two ORRs is the sum of the group delay responses of the individual ORR. In this way, the maximum delay and the bandwidth of the response can be increased.

3.2 Optical beamforming chip

A full photonic BFN is obtained by combining the ORR based delay elements with power splitters and combiners. An example of a 16×1 photonic BFN is shown in Fig. 4. It is based on a binary tree topology, consisting of four sections, 16 inputs and one output (Meijerink et al., 2010). In this particular case a total of twenty rings are involved. The rationale for using such a topology is that, for a linear PAA, increasing delay tuning ranges are required for the sixteen possible paths through the photonic BFN. Based on the photonic BFN architecture presented here, sufficiently large delay over a wide bandwidth has been demonstrated. A delay as high as 1.2 ns (for a comparison, 1 ns is approximately 30 cm of propagation distance in vacuum) over a bandwidth of 2.5 GHz has been demonstrated with a cascade of 7 ORRs, in a 8x1 optical beamformer (Zhuang et al., 2007). The details of the photonic BFN design can be found in (Meijerink et al., 2010, Zhuang et al., 2010).

In the photonic BFN, the signal processing (i.e. delaying and combining) is done in the optical domain. For this reason, the received RF signals from the AEs need to be converted from the electrical domain to the optical domain. This is done using optical modulation in the optical modulators. As shown in Fig. 4, an array of 16 optical modulators interface with the 16x1 photonic BFN chip. The signal flow in this photonic BFN is shown in detail. The optical power from a continuous wave (CW) laser is injected to the input of the optical splitter chip (Fig. 4a). A tunable coupler is used to split the optical power into two paths: one path goes to a 1x16 splitter and then to the optical modulators (in this case, Mach-Zehnder modulators (MZMs)), while the other goes to an unmodulated path, which later on will be used for the carrier re-insertion. Meanwhile the received signal from the AEs are amplified with low noise amplifiers (LNAs) and supplied to the RF input of the MZMs (Fig. 4b). The MZMs are biased at the minimum transmission point, creating a double sideband-suppressed carrier modulated optical signals (Fig. 4c). Next, the ORRs in the photonic BFN are tuned to provide the desired delay to one of the signal sidebands (Fig. 4d). The unwanted sideband then is filtered using and optical sideband filter (OSBF) (Fig. 4e). This yield an optical single sideband suppressed carrier (OSSB-SC) signal. In order to restore the modulating RF signal, the OSSB-SC signal has to be combined with the re-inserted optical carrier in the carrier re-insertion coupler (Fig. 4f). The signal containing the desired sideband and the optical carrier is then detected using a balanced photodetector (BPD) (Fig. 4g). This detection scheme allows cancellation of common noise and distortion terms contained at the two outputs of the re-insertion coupler.
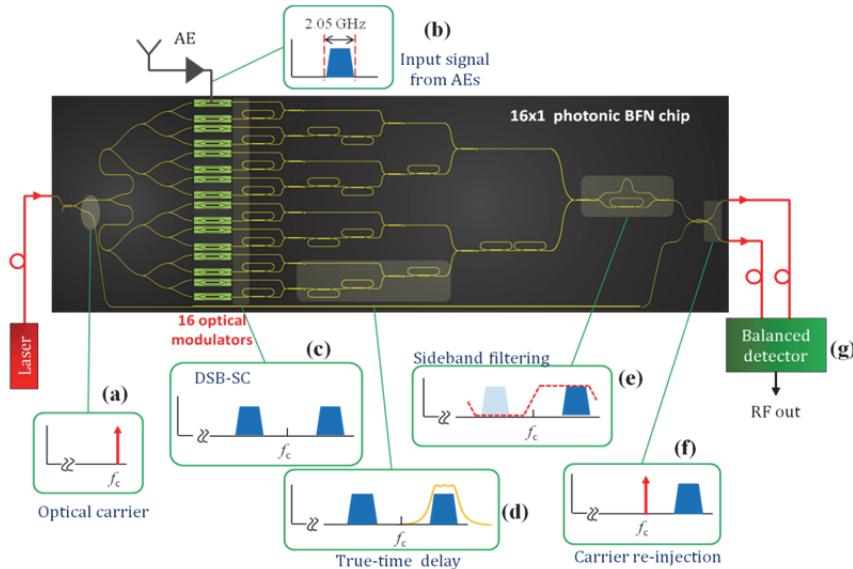
Fig. 4. Principle of operation of the photonic BFN system. The optical carrier from a laser (a) is modulated by the RF signals from the AEs (b) in the optical modulator array, yielding a DSB-SC spectrum (c). One of the sidebands of the optical signals are delayed and combined in the BFN (d) while the other is removed using an optical filter (e). The optical carrier is then recombined with the desired signal sideband (f) prior to the photodetection process in the balanced photodetector (g).

4. Phased array antenna system design

The photonic BFN explained in the previous section is the core of the total PAA system developed in this work. However, to beamform the signals from as many as 2048 AEs, one cannot use only a single photonic BFN. For this reason a scheme with two stages of optical beamforming is considered here (Marpaung et al., 2011, Meijerink et al., 2010). Two possibilities of how such a scheme can be implemented are illustrated in Fig. 5. In the first option (Fig. 5a), a tile consisting of 64 AEs is beamformed using a 64x1 photonic BFN. The (RF) outputs of thirty-two of these 64x1 photonic BFNs are then beamformed in the second stage using a 32x1 photonic BFN. This is the most straightforward implementation of the cascaded stages of photonic BFNs. However from previous investigations, it was concluded that a 64x1 photonic BFN presents a high degree of complexity and in turn imposes a higher risk in system reliability.

A way to avoid using very large photonic BFNs in the first stage is to combine RF beamforming and optical beamforming schemes as shown in Fig. 5b. In this scheme, every group of 4 AEs are combined using an RF beamforming scheme. Here, either (RF) time delay or phase shifting can be implemented. The latter will induce a beam squint but the effect is negligible since the required time delay between the neighboring AEs is relatively small (in the order of 34 ps). The use of RF beamforming will reduce the size of the photonic BFNs in the first stage to 16x1 instead of 64x1. These 16x1 photonic BFNs are well-investigated and relatively more reliable. Moreover, the scheme allows the use of an array of

16 modulators to interface with the photonic BFN instead of an array with 64 modulators which increases the yield of fabrication.
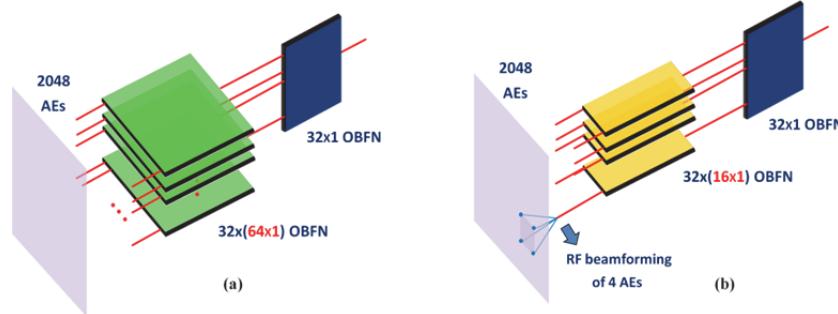
Fig. 5. Two options of implementing the two-stage optical beamforming scheme for a PAA with 2048 AEs. (a) In this scheme every 64 AEs in a tile is beamformed by a 64x1 photonic BFN. Outputs from 32 of these BFNs are combined in the second stage by a 32x1 photonic BFN. (b) A similar principle as in (a) but every group of 4 AEs is combined using RF beamformer. The first optical beamforming stage now consists of 32 of 16x1 photonic BFN.
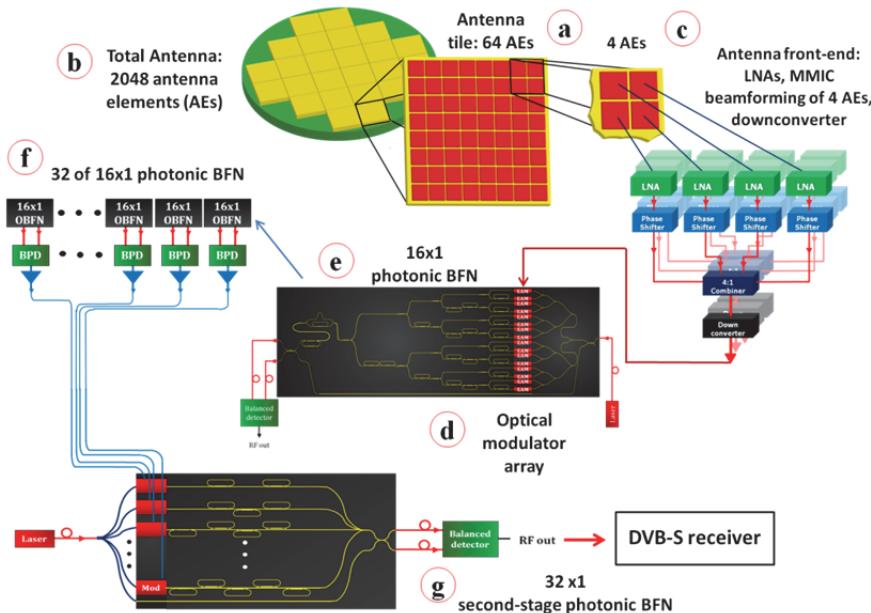
Fig. 6. The phased-array antenna system considered in this work. (a) An antenna tile consisting of 64 elements, (b) the total antenna system with >1800 AEs, (c) the RF front-end consisting of LNAs, MMIC phase shifter, a 4-to-1 combiner and a downconverter, (d) an array of optical modulator integrated with a 16x1 optical beamformer (e). The RF outputs of 32 of these 16x1 beamformer (f) are combined by a second stage optical beamfomer with 32 inputs and one output (g). BPD: balanced photodetector.

A more detailed schematic where the scheme in Fig. 5b is implemented is depicted in Fig. 6. The antenna tile (an 8x8 module) basic building block and the total antenna are shown in Figs. 6a and 6b, respectively. The antenna elements are stacked patch antennas that are impedance matched between 10.7 and 12.75 GHz (Verpoorte et al., 2011). In the antenna front-end, a chain of low noise amplifier (LNA) and monolithic microwave integrated circuit (MMIC) tunable phase shifter follows each AE. Using these phase shifters and a combiner, the RF beamforming is applied to a sub-array consisting of 4 neighbouring AEs. The output signal from the combiner is then down-converted to the L-band (Fig. 6c). The signals from the front end of each tile are then fed to an array of 16 optical modulators, to convert these signals to the optical domain (Fig. 6d). Each modulator array will interface with a 16x1 photonic BFN (Fig. 6e). Then, a larger photonic BFN with 32 input ports is used at the second-level to combine the outputs of 32 16x1 BFNs in the first stage (Fig. 6f and 6g). In the following section, the modelling and the performance evaluation of the entire system depicted in Fig. 6 is presented. The aim of this model is to retrieve the required system parameters of each component to meet the target specifications listed in Table 1.

4.1 System modeling

A procedure to simplify the complex PAA system depicted in Fig. 6 has been developed in order to analyze the system performance. The detail on the simplification procedure is beyond the scope of this chapter and has been reported elsewhere (Meijerink et al., 2010). The main idea is to reduce the dimension of such PAA system with multiple inputs and a single output into a two-port cascaded system. Eventually, to analyze such system the standard Friis' formula can be implemented (Meijerink et al., 2010). The resulting two-port model with its relevant parameters is depicted in Fig. 7. The system is now reduced to a cascade of an equivalent antenna and two blocks of receivers, each comprising an amplifier (the front end in the first block and the second-stage amplifier in the second block) and a two port model of the photonic BFNs (16x1 BFN in the first block and 32x1 BFN in the second block).
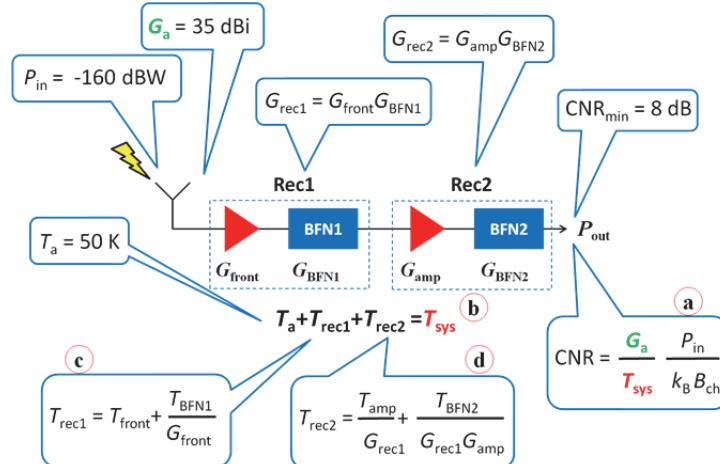

Fig. 7. Simplified two-port model of the PAA system in Fig.6.

For the system level simulation, some of the system parameters need to be fixed. From the antenna side, the gain and the antenna temperature are set to be 35 dBi and 50 K,

respectively. The antenna itself should receive signals from satellites, with a power level in the order of -150 dBW (assuming a satellite equivalent isotropic radiated power (EIRP) of 51 dBW) per transponder of 33 MHz bandwidth (Morello & Mignone, 2006). As will be shown later on, for some scenarios a margin up to 10 dB has been considered for this input power, leading to an input power as low as -160 dBW. The other parameters used in the simulation are listed in Table 2.

The model depicted in Fig. 7 shows the evolution of the gain and the noise temperature throughout the system. At the output, the system performance is evaluated in terms of the carrier to noise ratio (CNR) in the transponder bandwidth that should amount to at least 8 dB. This CNR is related to the G/T parameter mentioned in an earlier section by the equation (a) shown on Fig. 7. Using the values for the Boltzmann constant k_B=1.38×10^{-23} and B_{ch}=33×10^6 Hz, the CNR in dB and the G/T in dB/K is related by

$$\mathrm{CNR}(\mathrm{dB}) = G / T_{\mathrm{sys}}(\mathrm{dB} / \mathrm{K}) + P_{\mathrm{in}}(\mathrm{dBW}) + 153.5. \tag{1}$$

As an example, to achieve an 8 dB minimum CNR with the antenna with G/T of 12.7 dB/K, as listed in Table 1, the minimum received power should amount to -158.2 dBW.

Having the model of the system established, the design objective now is to maximize the CNR in equation (a) of Fig. 7. To do so, the system noise temperature (T_{sys}) should be minimized, since in this study the other parameters (P_{in}, G_a and B_{ch} are fixed). As shown in equation (b) of Fig. 7, T_{sys} comprises the noise temperatures of the antenna (T_a) and the two equivalent receiver blocks (T_{rec1} and T_{rec2}). Again we consider a fixed value of T_a, thus dictating that T_{rec1} and T_{rec2} should be minimized.

Simulation parameter	Symbol	Value
Antenna noise temperature	T_a	50 K
Received signal power	P_{in}	Varied: -160 dBW to -150 dBW
Antenna gain	G_a	35 dBi
Transponder bandwidth	B_{ch}	33 MHz
Minimum output carrier-to-noise ratio in the transponder bandwidth	$\mathrm{CNR}_{\min}$	8 dB
Laser optical power	P_{laser}	100 mW
Laser relative intensity noise (RIN)	RIN	-150 dB/Hz
Balanced photodetector (BPD) responsivity	r_{PD}	0.8 A/W
BPD common mode rejection ratio	CMRR	30 dB
Load impedance	R_L	50 ohm
Second stage amplification gain	G_{amp}	30 dB
Second stage amplification noise figure	$\mathrm{NF}_{\mathrm{amp}}$	3 dB
Fiber-to-chip coupling loss	L_f	1 dB
Passband loss of the optical sideband filter	L_{OSBF}	1 dB
Front-end gain	G_{front}	Target: 70 dB
Front –end noise figure	$\mathrm{NF}_{\mathrm{front}}$	Target: 2.5 dB
Modulator half-wave voltage	V_π	Target: 4 V
Modulator insertion loss	L_m	Target: 5 dB
Optical waveguide propagation loss	L_{wg}	Target: 0.2 dB/cm

Table 2. System parameters used in the simulation and the performance analysis.

As shown in Friis' formulas in (c) and (d) of Fig. 7, these noise temperatures depend on the properties of the amplifiers and the BFNs. To minimize T_{rec1} one has to use a front-end with high gain and low noise figure (NF). Moreover, the noise temperature contribution of the first stage BFN (T_{BFN1}) should be kept low, such that at the end $T_{rec1} \approx T_{front}$. The BFN noise temperature itself is inversely proportional to the BFN gain (G_{BFN}) i.e.

$$T_{\text{BFN}} \propto \frac{1}{G_{\text{BFN}}}. \tag{2}$$

Thus, maximizing the gain of the BFN is paramount in achieving the required CNR. The BFN gain depends on the parameters of the optical components in the system, namely the laser, photodetector, the optical modulators and the photonic BFN chip. In this study we focus on the dependence of G_{BFN} to three key parameters, the optical waveguide propagation loss (L_{wg}) and the modulator half-wave voltage (V_π) and insertion loss (L_m). As shown in Eq. 3, G_{BFN} is inversely proportional to the square of the aforementioned parameters.

$$G_{\text{BFN}} \propto \frac{1}{\left(V_\pi \, L_{\text{m}} \, L_{\text{wg}}\right)^2} \tag{3}$$

Note that here the loss is defined in the range of $[1, \infty]$ such that $L=1$ indicates a lossless system. Thus, it is important to limit the propagation loss in the optical waveguides as well as to ensure a highly efficient modulation in the modulators, characterized by low half-wave voltage and insertion loss. The optimum values of these parameters, together with the front-end gain and noise figure, will be determined from the system simulations in order to achieve the CNR_{min} specification. The other system parameters used in the simulation are also listed in Table 2.

4.2 System performance

In Fig. 8, the system CNR is shown as function of the optical waveguide loss in dB/cm at three different conditions. First, the front-end gain and the modulator V_π are set at 70 dB and 1 V, respectively. The rest of the parameters used are: P_{in}=-160 dBW, L_m=2 dB and NF_{front} = 0.7 dB. This, in fact, is describing the configuration with very high performance modulators and front-ends. With these parameters, the 8 dB CNR can be met for all waveguide loss values from 0 to 0.5 dB/cm (solid curve). But if the V_π increases to 4.0 V (which is more commonly found in commercial off-the shelf modulators), the specified CNR can only be met with L_{wg} of 0.2 dB/cm or lower (dashed-curve). If now the front-end gain is reduced to 60 dB, the CNR specification cannot be met with even with lossless optical waveguides (dash-dotted curve) (Marpaung et al., 2011). From these simulation results, the target value for G_{front} is set to 70 dB while the target value for the maximum L_{wg} is 0.2 dB/cm.

With a front-end gain of 70 dB and a low waveguide propagation loss the system noise temperature is likely to be dominated by the noise temperature of the front-end, T_{front} which is related to the front end noise figure (NF_{front}) via the relation:

$$\text{NF}_{\text{front}} = 10 \log_{10}\left(1 + \frac{T_{\text{front}}}{290}\right). \tag{4}$$
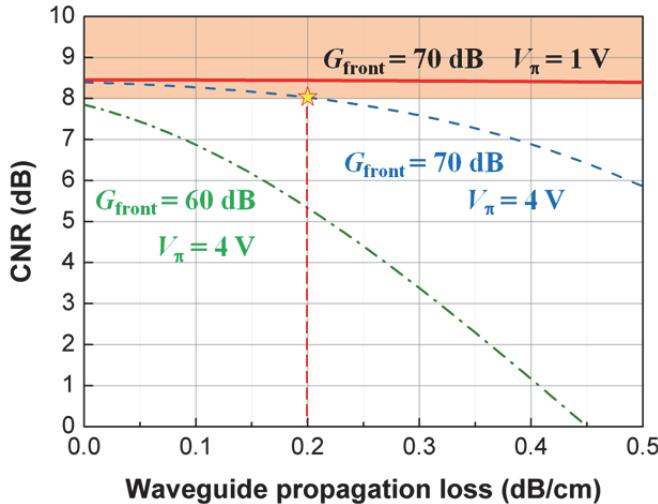
Fig. 8. Simulation results used to determine the requirements for G_{front} and L_{wg}. The system CNR is depicted as function of L_{wg} where G_{front} and V_π are used as parameters. The target L_{wg} value of 0.2 dB/cm is indicated by a star. The input signal power is -160 dBW.

In the simulation result presented in Fig. 8, it was assumed that T_{front} = 50 K what corresponds to NF_{front} = 0.7 dB. It could be challenging to achieve such a low value of NF. Thus, a simulation was carried out to determine the maximum front end NF that can be tolerated by the system to still achieve the target CNR. The result is depicted in Fig. 9.
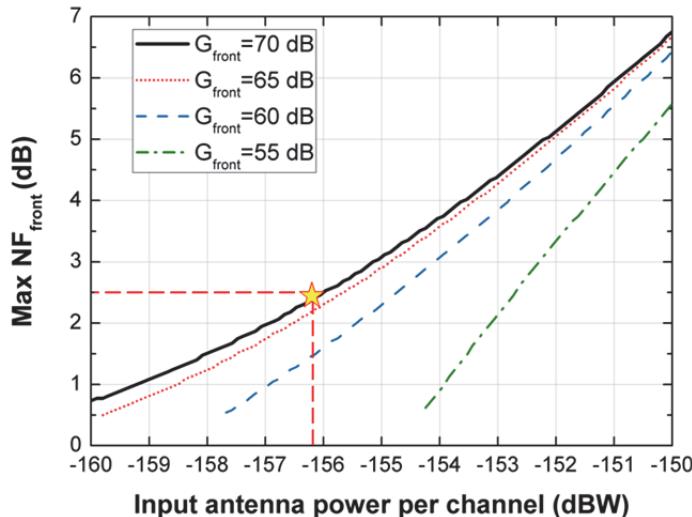

Fig. 9. Simulation results used to determine the maximum allowed NF_{front} as a function of the received signal power. The parameter values used are: V_π = 4 V, L_m = 2 dB and L_{wg} = 0.2 dB/cm. A target value of 2.5 dB front-end NF has been set, as indicated by a star symbol.

From the figure above one can determine that in order to achieve the required CNR with a received power of -160 dBW, the front end gain should be > 60 dB and NF < 1 dB. Since this is very challenging to achieve, the system requirement aimed in this work is slightly relaxed such that a power margin of 6 dB (input power of -156 dBW) is sufficient. Thus in this case, the NF requirement of the front-end is relaxed to < 2.5 dB for gain of 70 dB. These will be the target values for the front-end design.

Finally, to determine the parameters of the optical modulators, the simulation results in Fig. 10 are used. Here, the maximum allowed V_π to achieve 8 dB CNR is depicted as a function of the received RF power, with the modulator insertion loss is kept as a parameter. Although the front-end gain and NF are slightly different from the one determined in the previous simulation, the modulator parameters can still be determined. As mentioned earlier, to achieve the value L_m = 2 dB and low V_π simultaneously, one should employ a very high performance modulator. As suggested from the result in Fig. 10, to relax these requirements one should sacrifice a margin in the received power. Here, we have set preliminary target values of L_m = 5 dB and V_π = 4 V for the modulators, which allow a minimum received power of -154 dBW (i.e. a 4 dB power margin).
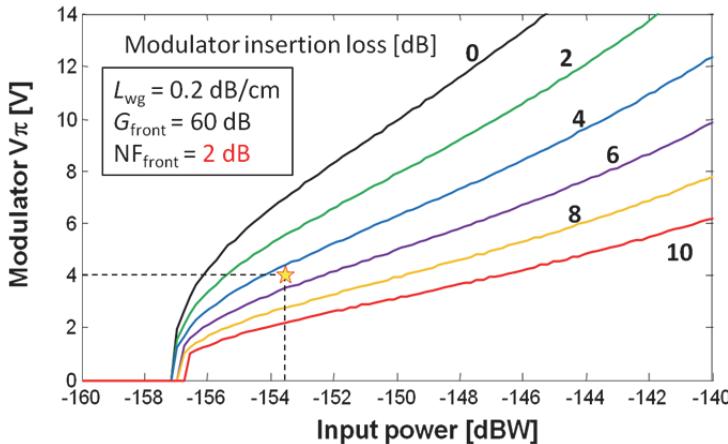
Fig. 10. Simulation results used to determine modulator parameters, V_π and L_m to achieve a minimum CNR of 8 dB. Preliminary target values of L_m = 5 dB and V_π = 4 V have been set, as indicated by a star symbol.

In the following section, the progress on the development of the components to achieve the target values determined earlier will be discussed.

5. Components development

From the system level simulation presented in the previous section the required characteristics of each component of the system have been specified. In the current phase of the SANDRA project numerous efforts have been directed towards meeting the specified performance. The development of three main parts that constitute an antenna tile (64 AEs), namely the front-ends (LNAs, phase shifters and downconverter chips), array of 16 optical modulators and the 16x1 photonic BFN chips are described in the following subsections.

5.1 Front-end

The projected configuration of the front-end is depicted in Fig. 11a. Here, the output signal of 4 AEs will be amplified, combined and subsequently downconverted to the L-band (950 3000 MHz). The target values of gain and noise figure for the complete chain have been set to 70 dB and 2.5 dB, respectively. From the front-end design point of view, downconversion to the L-band is advantageous to avoid oscillations due to the large amplification. The oscillations can be minimized by means of distributing the gain at the two frequency bands.

Due to size constraints it is desirable to use MMICs with a high degree of integration. For this reason the so-called corechips will be used in the design. These are MMIC-building blocks that integrate different functionalities in the same chip. Here, the combined functionalities of amplification (LNA) and phase shifting are desired. The one selected for this design is a corechip that was previously designed for the NATALIA-project (Baggen et al., 2010) and manufactured by the foundry OMMIC. It consists of a two-stage LNA, 4 bit phase shifter and digital logic. The projected gain and NF of the corechip are 12 dB and 1.7 dB, respectively, with an assumption that the coupling loss from the antenna to the chip is less than 0.5 dB.

| OVERVIEW OF GAIN AND NF			
Element	Gain	NF	Components functionality
Corechip	12	1.7	Amplification and Phase Shifting
4-to-1 Combiner	-1	1	Constructing a 2x2 Subgroup
Ku LNA	13	2	Second Stage Amplification
Image Filter	-3	2	Rejecting Image Frequencies
NXP Chip	35	9	Down-conversion and Amplification in Ku and L band
L-band Amplifier	14	4	Gain Compensation in L band
Total System	70	2.4	

Fig. 11. (a) The configuration of the front-end. The corechip consists of an LNA and a 4-bit phase shifter. The NXP chip is used for downconversion and amplification. (b) Schematic layout of the RF front-end. (c) Overview of the gain and noise figure of the elements of the front-end chain.

The outputs of the four corechips are subsequently combined in a 4:1 combiner followed by an LNA. The down-conversion is performed after combining the 2x2 sub array. The proposed chip for down-conversion here is manufactured by NXP. This chip is a highly integrated circuit that includes an LNA, a mixer, a down-converter, a phase-lock loop, a crystal oscillator, and an intermediate-frequency buffer. This chip supports RF input frequencies between 10.7 and 12.75 GHz, and uses a selectable LO that operates at 9.75 or

10.6 GHz. Finally an L-band amplifier is placed after the downconverter to achieve the gain in the L-band. The preliminary front-end layout of a 2x2 sub-array is depicted in Fig. 11b. The overview of the projected gain and noise figure of the complete front-end chain is summarized in Fig. 11c. A gain of 70 dB and a noise figure of 2.4 dB are achievable with this design which is in line with the target values set by the system simulations.

5.2 Optical modulator array

The electro-optic modulators developed in this work are surface-normal electroabsorption modulators (EAMs) based on InGaAs/InAlAs coupled quantum wells (Q. Wang et al., 2006). Compared to traditional waveguide EAMs, surface-normal EAMs offer significant advantages in terms of polarisation insensitivity, large active apertures and low insertion losses. A drawback of these modulators is the short interaction length between the incident light and the active medium, thus limiting the contrast ratio. Single-path surface-normal EAMs have typical contrast ratios in the range 2:1.
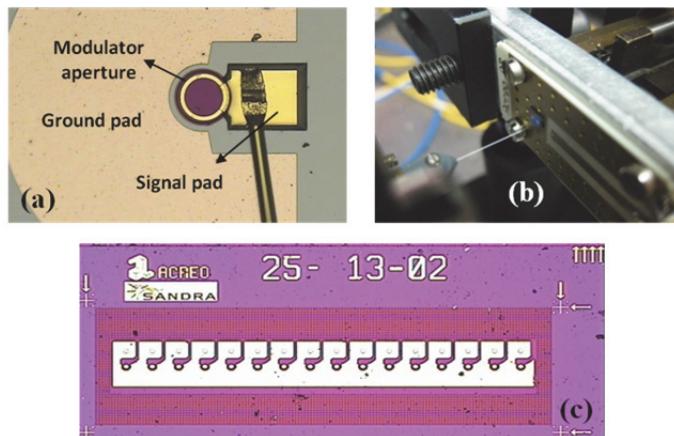

Fig. 12. (a) A microscope image of a surface normal electroabsorption modulator (EAM). (b) A photograph of the transmissive EAM mounted on a PCB. (c) A microscope image of a reflective EAM array consisting of 16 modulators. The pitch of the array is 127 μm and the chip size is 3.1 mm x 1.5 mm

The structure of a single modulator is shown in Fig. 12a, depicting a large area of ground pad, the modulator circular aperture and a pad to supply the reverse bias and the RF signal voltages. In this work, two types of EAMs with different diameters, ranging from 125 μm down to 25 μm, were fabricated. The first EAM type is the transmissive (i.e. single-path) modulator and the second type is the reflective (i.e. double-pass) modulator. The reflective EAM is obtained by means of depositing a mirror on one side of the modulator while the other side is anti-reflection coated. Thus, the incident light will experience twice absorption in the active area and the modulation contrast ratio will be enhanced but at the expense of an increased insertion loss. For test purposes, the modulator is mounted on a PCB using wire bonds. An SMA connector is then soldered to the PCB to supply the required reverse bias voltage and the RF signal. A photograph of a transmissive modulator on the PCB with optical fibers coupled at the input and output is shown in Fig. 12b. In Fig. 12c, the

photograph of an array of 16 reflective EAMs is depicted. The aperture size of these modulators is 25 μm. At the final system, this type of array will be interfaced with photonic the BFN chip by means of hybrid integration.

5.2.1 DC characterization

The characterization of the EAM started with the DC characteristics measurements. A tunable laser diode (TLD, Santec) is used to supply the light to the modulator and the output intensity is measured with an optical power meter (HP 8153A). Meanwhile the reverse bias voltage of the modulator is controlled using a variable power supply (HP 6634A). Various measurements were performed where all of them were automated in LabVIEW. First, the optical wavelength was swept from 1520 nm to 1580 nm with a step of 1 nm and the bias voltage was kept as a parameter and varied from 0 to 12 V with a step of 1 V. The optical power from the TLD was kept at 0 dBm. The transmission over the EAM as functions of the wavelength, with varying bias voltage, is shown in Fig. 13a. These results show that the EAM is more sensitive to bias voltage variations at two regions, roughly from 1530 nm to 1550 nm and at 1555 nm to 1570 nm. The modulator behaviours at these regions are different. In the lower wavelength region, an increase in the bias voltage results in a decrease in the transmission (or an increase in the absorption). In contrast, at the higher wavelength region, an increase in bias voltage results in an increase of transmitted optical power.

We also measured the transmission as a function of the bias voltage, for several optical wavelength values. This measurement is important to inspect the static linearity of the modulator transfer function (transmission vs. bias voltage). The results are shown in Fig. 13b, where a range from 1520 nm to 1545 nm is considered. We can see relatively linear responses are obtained in the bias voltage region of 2-8 V for a wavelength region of 1539 nm to 1541 nm. For λ = 1545 nm, the static characteristic is relatively linear up to 12 V of bias voltage. In the next parts, the RF characterization results of the modulators are reported.

(a)

(b)
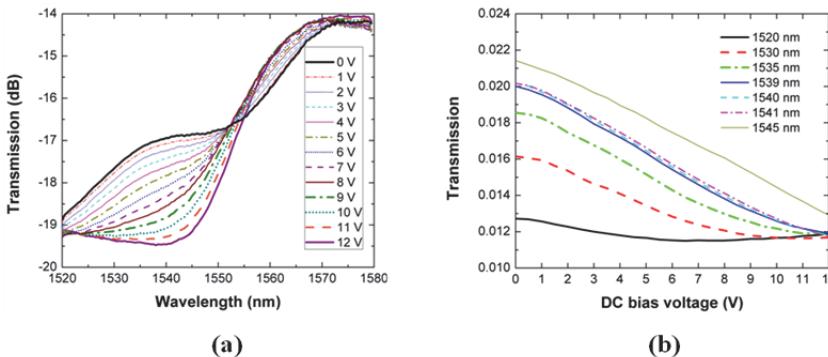
Fig. 13. DC characterisation results on a single pass (transmissive) EAM with an aperture of 100 μm. The input optical power to the modulator is set at 0 dBm. (a) The transmission through the modulator (in decibels) as a function of the optical wavelength for different reverse bias voltages. (b) Transmission (in linear scale) as a function of the reverse bias voltage, for various optical wavelengths from 1520 nm to 1545 nm

5.2.2 RF characterization: speed

The modulator speed depends on the size of the modulator. For a rough estimation of its bandwidth, the modulator can be modelled as an RC circuit where the capacitance depends linearly on the aperture area of the modulator. Thus, the smaller the modulator aperture, the higher cut-off frequency will be. For a relatively small EAM with an aperture diameter of 25 μm, the 3 dB cut-off frequency is expected to be in the order of 12 GHz. This is well above the intended frequency range of operation in the L-band (950-3000 MHz) resulting from the downconversion in the front-end.

To verify the modulator bandwidth, a measurement on the modulator frequency response (s_{21}) is carried out. Here the device under consideration is a reflective EAM with an aperture size of 25 μm. The measurement result for various reverse-bias voltages is depicted in Fig. 14a. Here, a laser with an optical wavelength of 1530 nm and an optical power of +9 dBm has been used as the optical source. It can be seen from the figure that in an extended frequency range of operation (1-5 GHz, marked as the shaded area in Fig. 14a) the modulator shows a relatively flat frequency response, with a 6-dB bandwidth of 5 GHz for a reverse bias of 3 V and 5 V (Marpaung et al., 2011).
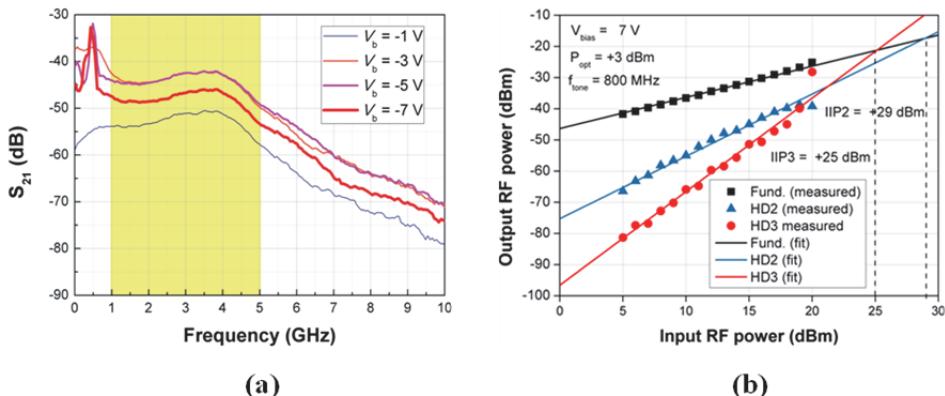

Fig. 14. RF characterization results on the EAMs. (a) The measured frequency response (s_{21}) of a reflective EAM with a 25 μm aperture for various reverse-bias voltages. The optical wavelength of the laser is 1530 nm. (b) The measured fundamental tone, HD2 and HD3 powers of a 100 μm aperture transmissive EAM for the bias voltage of 7 V at 1545 nm.

It is important to point out from Fig. 14a that the magnitude of the frequency response is relatively low (approximately -45 dB). The origin of this large RF-to-RF loss is still under investigation. Currently a lot of efforts are towards the improvement of the design and the quality of the wirebonds and the RF PCB used to mount the modulator. These improvements expectedly will lead to an accurate determination of the modulator V_π.

5.2.3 RF characterization: nonlinearity

Preliminary nonlinearity measurements have been performed on the transmissive EAM. For the particular EAM used in the experiments (100 μm aperture) the cut-off frequency is in the order of 1 GHz. A single-tone measurement was performed to probe the nonlinearities in the EAM. A modulating tone with a frequency of 800 MHz was supplied using an RF signal

generator. The EAM was reverse biased at 7 V while the optical wavelength was chosen at 1545 nm. The optical power from the TLD was set at +3 dBm. The RF tone power was then swept from +5 dBm up to +20 dBm with a step of 1 dB. The photodetector RF power was then measured with an RF spectrum analyzer at the fundamental, second-order harmonic (HD2) and third-order harmonic (HD3) distortions frequencies of 800 MHz, 1.6 GHz and 2.4 GHz, respectively. The measurement results are depicted in Fig. 14b. From these measurements, the 2nd-order and 3rd-order input intercept points (IIP2 and IIP3) of the EAM are determined to be +29 dBm and +25 dBm, respectively. As a comparison, for a Mach-Zehnder modulator with $V_\pi = 4$ V, the IIP3 is +21 dBm (Marpaung et al., 2010). Hence, the EAM under test under test have shown lower third-order nonlinearity relative to the aforementioned MZM. Currently, we are in the stage of extending these RF measurements to both types of modulators (transmissive and reflective) for various sizes.

5.3 16x1 photonic beamfomer chip

The photonic beamformer chip is developed using TriPleX waveguide technology (Bauters et al., 2011, Morichetti et al., 2007) that allows both low propagation loss and small bending radius to be achieved simultaneously. Three aspects regarding the 16x1 photonic BFN chip discussed here are the waveguide propagation loss, the layout of the photonic chip and the design of the optical sideband filter.

5.3.1 Waveguide propagation loss

As mentioned in the previous section, the target value for the maximum waveguide propagation loss is 0.2 dB/cm (at the optical wavelength range of 1530-1570 nm). Various test structures (for example directional couplers, spirals and ORRs) have been developed in the TriPleX technology with an optical waveguide structure consisting of a double stripe of which the cross-section is shown in the inset of Fig. 15.
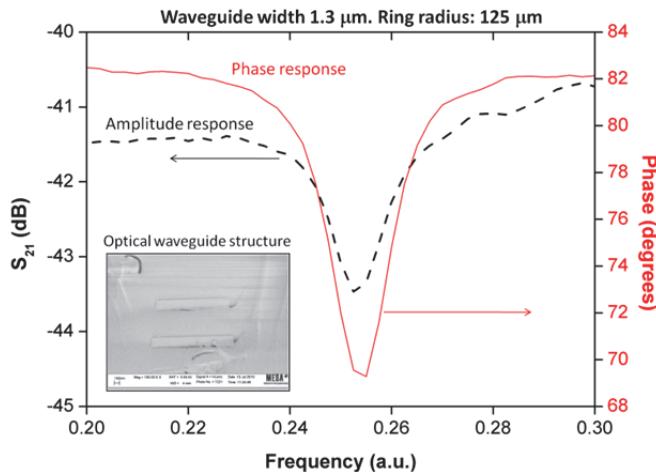

Fig. 15. Result of the propagation loss characterization of an optical ring resonator using the phase shift method. A loss of 0.2 dB/cm has been achieved. Inset: A scanning electron microscope (SEM) image of the waveguide cross-section.

A propagation loss measurement was performed in an ORR structure using the phase-shift method reported in (Roeloffzen et al., 2005). In this method, the ring resonators are tuned using a heater controller to its resonance frequency. The magnitude and the phase responses of this ring are then measured. The results are shown in Fig. 15. The measurement was performed on an ORR with a waveguide width of 1.3 μm and radius of 125 μm. This particular value was chosen to ensure that the measured loss is dominated by the waveguide propagation loss instead of the bend loss. From these measurements the propagation loss of the optical waveguide can be estimated to be as low as 0.2 dB/cm for TE polarized light which means that the target value from the system simulations has been met. Furthermore, from the simulation results it is expected that the bend loss will not become the dominant factor for a bending radius as low as 75 μm. It is important to mention that at these small bending radii a signifcant reduction of the optical beamformer chip can be realized as compared to previously developed BFNs (Marpaung et al., 2011).

5.3.2 Photonic chip layout

The 16x1 photonic BFN chip is designed to meet the criteria listed in Table 1. The layout of such a BFN follows the previous designs which use binary tree architecture. The important step in the design is to determine the optimum number of ORRs used in the chip. Due to the time-bandwidth product limitation explained in Section 3, the number of ORRs involved is estimated from the required maximum time delay and the signal bandwidth.

The maximum time delay in the BFN can be estimated from the information of the scanning angle, inter-element distance and how these elements are arranged in the tile. Fig. 16a shows the schematic of an 8x8 tile of 64 AEs. Here, d_{AE} is the distance between the antenna elements. Since an RF beamforming scheme is implemented in every group of 2x2 elements, the photonic BFN only "sees" the elements marked in (dark) red in Fig. 16a. The distance between the neighboring elements seen by the photonic BFN in this case is $d_{BFN}=2d_{AE}$. It can then be calculated that the maximum time delay between the elements seen by the 16x1 photonic BFN in this arrangement is

$$t_{max} = 3\sqrt{2}\ t_{BFN} = 6\sqrt{2}\ t_{AE} \tag{5}$$

The time delay needed between the adjacent elements (t_{AE}) is related to d_{AE} as follows

$$t_{AE} = \frac{d_{AE}\sin\theta}{c_0}. \tag{6}$$

Here θ is the maximum elevation scanning angle and c_0 is the speed of light in vacuum. Using the value of $\theta = 60°$ and $d_{AE} = 1.18$ cm as listed in Table 1, one can calculate from Eqs. (5) and (6) that $t_{AE} = 34$ ps and subsequently $t_{max} = 290$ ps.

Although in this work the considered signal bandwidth is in the order of 2.05 GHz, the 16x1 photonic BFN is designed for larger bandwidth. The reason for this is to have the flexibility for the case that the antenna system needs to accommodate both the horizontal and vertical polarizations of the satellite signals in the future. Thus, in this case the minimum bandwidth for the BFN becomes 2x2.05 = 4.1 GHz. For this purpose the chip is designed to cover a bandwidth in excess of 4.3 GHz, which include a guard band.

It has been calculated that the time delay and bandwidth requirements derived earlier can be achieved with a BFN consisting of 40 ORRs. The resulting functional design and the

optical chip layout of the 16x1 BFN are shown in Fig. 16b and 16c, respectively. The BFN has been designed to be able to interface either with the transmissive or the reflective electroabsorption modulator arrays. A picture of the realized 16x1 BFN chip is depicted on Fig. 16d, together with a 20 cent Euro coin for size comparison. The total chip dimension of the BFN chip is 0.7 cm x 2.2 cm. This features a size reduction nearly 10 times compared to a 16x1 photonic BFN chip with a less complexity reported previously (Burla et al., 2010, Zhuang et al., 2010).
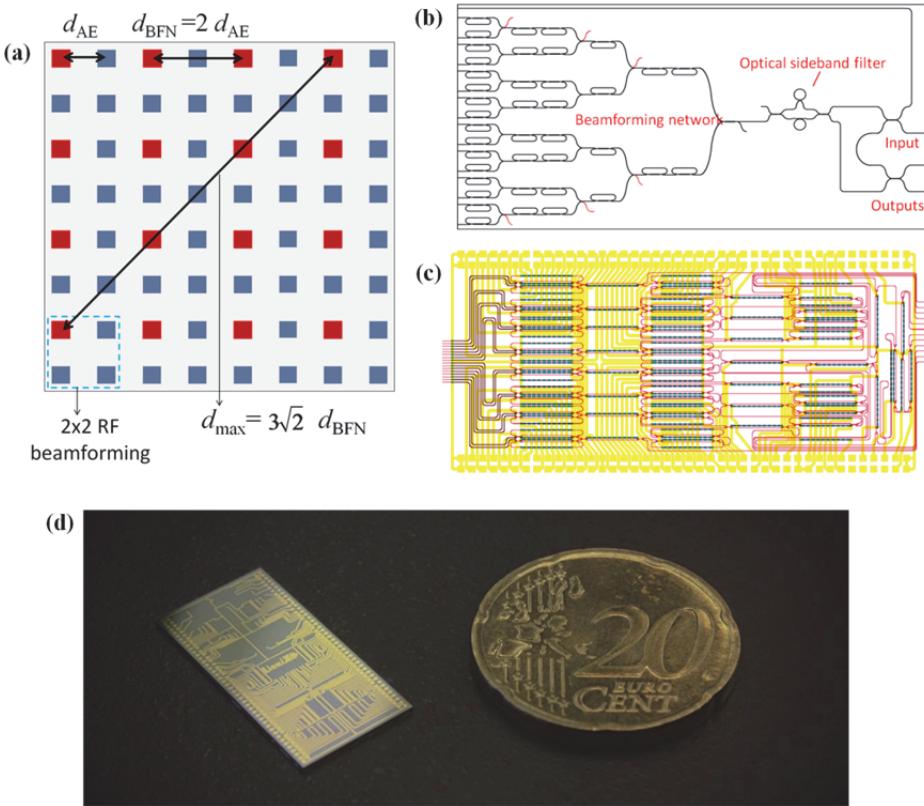
Fig. 16. (a) An antenna tile consisting of 64 AEs. (b) Functional design of the 16x1 photonic BFN showing the ORR delay elements and the sideband filter. (b) Chip layout of the BFN showing the optical waveguides, the heaters layout and the electrical wiring. The chip dimension is 0.7 cm x 2.2 cm. (d) The 16x1 photonic BFN chip pictured with a 20 cent Euro coin for size comparison.

5.3.3 Optical sideband filter

As mentioned earlier, the photonic BFN employs an OSSB-SC modulation scheme. In previous investigations (Meijerink et al., 2010, Zhuang et al., 2010), where MZMs instead of EAMs have been used, optical carrier suppression can be achieved by low-biasing the MZMs, while an optical filter is used to remove one of the signal sidebands. In that case a

Mach-Zehnder interferometer (MZI) with an ORR in one of its arms (MZI+1 ring) is used for the sideband filtering (Meijerink et al., 2010, Zhuang et al., 2010). In this work however, EA intensity modulators with a double-sideband with full carrier output spectrum are used instead of MZMs. Hence, an optical filter is required to suppress both the optical carrier and one of the sidebands. It turns out that an MZI+ 1 ring structure does not feature a transition that is sharp enough to do this. This is depicted in Fig. 17, where the measured and simulated responses of this filter are depicted, together with the position of the optical carrier. To improve the selectivity, an MZI structure where both arms are loaded with ORRs (MZI+2 rings) (Z. Wang et al., 2007) will be used for the filtering. The simulated response of such a filter is also depicted in Fig. 17, clearly depicting an improved selectivity and a narrower transition band. Both filters have been realized using the TriPleX waveguide technology. The waveguide layouts of these filters are depicted in Fig. 17. By means of fitting the measured response of the MZI+1 ring filter (Fig. 17), a waveguide propagation loss of 0.2 dB/cm is verified. The measurement on the MZI+2 rings filter is currently ongoing and will be reported elsewhere.
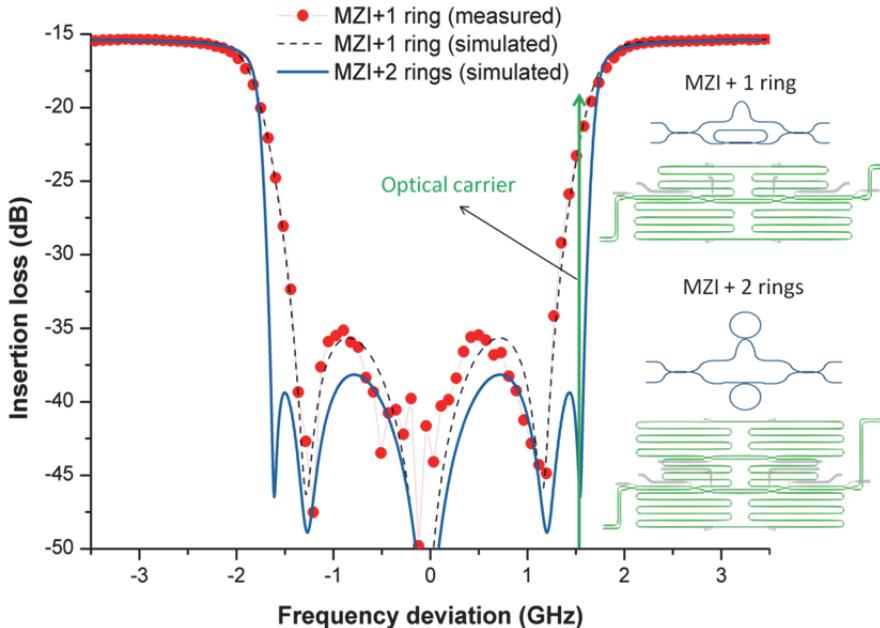
Fig. 17. Optical sideband filters measured and simulated responses. In the fitting a waveguide propagation loss of 0.2 dB/cm is used.

6. Photonic integration scheme

As mentioned in Subsection 3.2 the photonic BFN system employs optical single-sideband suppressed-carrier (OSSB-SC) modulation and coherent optical detection techniques. In this scheme to achieve proper combination of the signals optical phase synchronization of each branch of the BFN is required (Meijerink et al., 2010). To maintain the optical phase stability

in the photonic BFN chip itself has been shown to be viable (Zhuang et al., 2010). However, in the demonstration previously reported fiber pigtailed commercial off-the-shelf optical modulators have been used. This leads to a poor stability of the system. Thus, to depart from proof-of-concept towards an implementation in an actual PAA system, an important aspect that must be addressed is the photonic integration of the BFN chip and the optical modulator array.

A possible scheme for the integration of the 16x1 photonic BFN chip and the reflective EAM array is shown in Fig. 18. A carrier chip for the modulator array (fabricated using a silicon substrate covered by a thin SiO_2 film) is used to provide mechanical strength to the EAM chip as well as acting as the fan-out of the electrical paths going to the EAMs. The EAM chip is then flip-chip bonded onto the carrier chip. Before interfacing with the BFN chip the hybridization of the modulator chip is required. In this step the InP substrate of the EAMs has to be thinned down to reduce the insertion loss between the BFN chip and the modulator. It can be calculated that a substrate thickness of below 10 μm is required to achieve an insertion loss of below 3 dB between these two chips. The photonic BFN chip itself will be mounted on a PCB to provide the electrical paths to the heaters for thermo-optical tuning. The fiber-to-chip couplings of the laser and detector to the BFN chip will be done with butt-coupling. This photonic module will then be interfaced with the PCB containing the front-ends and the antenna tile using a connector array or a flex-cable.
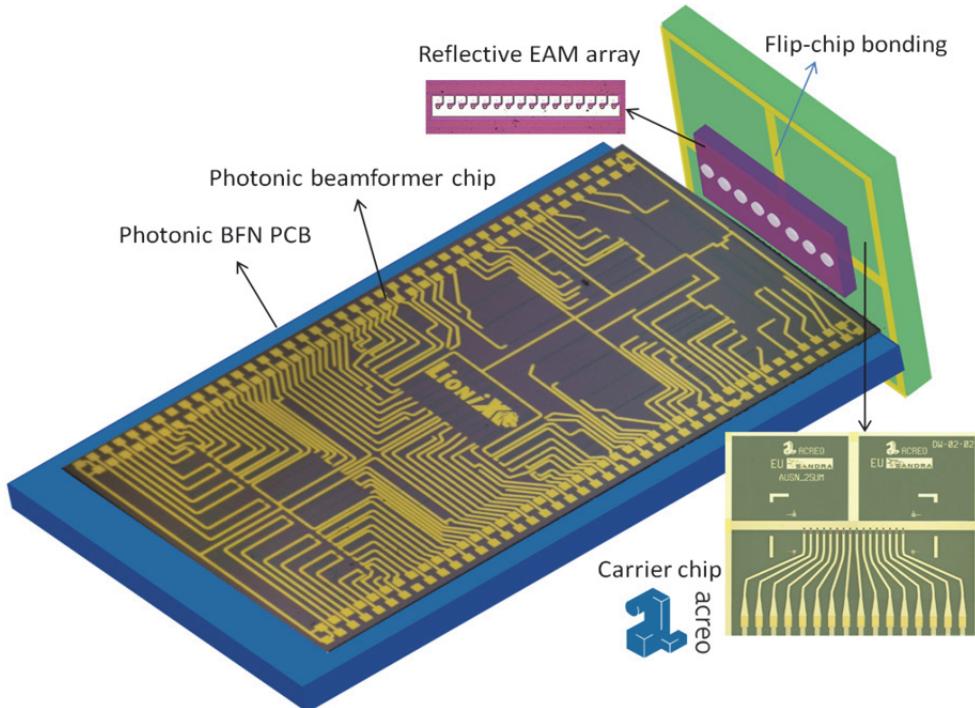

Fig. 18. An artist impression of a possible photonic integration scheme between the photonic BFN chip and the EAM array.

7. Conclusions

We have reported the design, performance analysis and the progress on components development of a novel Ku-band phased-array antenna system for airborne applications. A system level simulation has been used to determine the target values for the key parameters of the system components. Various target values like the front-end gain and noise figure as well as the propagation loss of the optical waveguide have been met. Development of two key components, namely the photonic BFN chip and the EAM array chip are reported. The first step towards the photonic integration of these chips is proposed.

8. Acknowledgment

9. References

Baggen, R.; Vaccaro, S.; del Rio, D. ; Sanchez, R. & Langgartner, G. (2010). First Prototyping of a Compact Mobile K_u-band Satellite Terminal. *Proceedings of the 4th European Conference on Antennas and Propagation (EuCAP 2010),* pp. 1-5, ISBN 978-84-7653-472-4, Barcelona, Spain, April 2010

Baggen, R.; Holzwarth, S.; Böttcher, M. & Sanadgol, B. (2011). Phased Array Technology for Mobile User Terminals. *Proceedings of the 5th European Conference on Antennas and Propagation (EuCAP 2011),* pp. 2782-2786, ISBN 978-88-8202-074-3, Rome, Italy, April 2011

Bauters, J.; Heck, M.; John, D.; Dai, D.; Tien, M.; Barton, J.; Leinse, A. ; Heideman, R.; Blumenthal, D. & Bowers, J. (2011). Ultra-Low-Loss High-Aspect-Ratio Si_3N_4 Waveguides. *Optics Express,* Vol. 19, No. 4, February 2011, pp. 3163-3174, ISSN 1094-4087

Burla, M.; Khan, M.R.H.; Marpaung, D.; Roeloffzen, C.; Maat, P.; Dijkstra, K.; Leinse, A.; Hoekman, M. & Heideman, R. (2010). Squint-free Beamsteering Demonstration using a Photonic Integrated Beamformer based on Optical Ring Resonators. *Proceedings of the IEEE Topical Meeting on Microwave Photonics (MWP 2010),* pp. 1-4, ISBN 978-1-4244-7824-8, Montreal, Canada, October 2010

Marpaung, D.; Roeloffzen, C.; Leinse, A. & Hoekman, M. (2010). A Photonic Chip based Frequency Discriminator for a High Performance Microwave Photonic Link. *Optics Express,* Vol. 18, No. 26, December 2010, pp. 27359-27370, ISSN 1094-4087

Marpaung, D.; Zhuang, L.; Burla, M.; Roeloffzen, C.; Verpoorte, J.; Schippers, H.; Hulzinga, A.; Jorna, P.; Beeker, W.P.; Leinse, A.; Heideman, R.; Noharet, B.; Wang, Q.; Sanadgol, B. & Baggen, R. (2011). Towards a Broadband and Squint-free K$_u$-band Phased Array Antenna System for Airborne Satellite Communications. *Proceedings of the 5th European Conference on Antennas and Propagation (EuCAP 2011),* pp. 2774-2778, ISBN 978-88-8202-074-3, Rome, Italy, April 2011

Meijerink, A.; Roeloffzen, C.; Meijerink, R.; Zhuang, L.; Marpaung, D.; Bentum, M. ; Burla, M. ; Verpoorte, J. ; Jorna, P. ; Hulzinga, A. & van Etten, W. (2010). Novel Ring Resonator-Based Integrated Photonic Beamformer for Broadband Phased Array Receive Antennas—Part I: Design and Performance Analysis. *Journal of Lightwave Technology,* Vol. 28, No. 1, January 2010, pp. 3-18, ISSN 0733-8724

Morello, A. & Mignone, V. (2006). DVB-S2 : The Second Generation Standard for Satellite Broad-Band Services. *Proceedings of the IEEE*, Vol. 94, No. 1, January 2006, pp. 210-227, ISSN 0018-9219

Morichetti, F.; Melloni, A. ; Martinelli, M.; Heideman, R.; Leinse, A.; Geuzebroek, D. & Borreman, A. (2007). Box-Shaped Dielectric Waveguides : A New Concept in Integrated Optics. *Journal of Lightwave Technology,* Vol. 25, No. 9, September 2007, pp. 2579-2589, ISSN 0733-8724

Riza, N.A & Thompson, J.B. (Eds.). (1997). *Selected Papers on Photonic Control Systems for Phased Array Antennas, Series SPIE Milestone Vol. MS136,* SPIE Press, ISBN 9780819426130, New York

Roeloffzen, C.; Zhuang, L.; Heideman, R.; Borreman A. & van Etten, W. (2005). Ring Resonator-Based Tunable Optical Delay Line in LPCVD Waveguide Technology. *Proceedings IEEE/LEOS Benelux Chapter 2005*, pp. 79-82, Mons, Belgium, December 2005

SANDRA project website, May 2011, Available from : www.sandra.aero

Verpoorte, J.; Schippers, H.; Jorna, P.; Hulzinga, A.; Roeloffzen, C.; Marpaung, D.; Sanadgol, B.; Baggen, R.; Wang, Q.; Noharet, B. ; Beeker, W.; Leinse, A. & Heideman, R. (2011). Development of the SANDRA Antenna for Airborne Satellite Communication. *Proceedings of the IEEE Aerospace Conference 2011,* pp. 1-15, ISBN 978-1-4244-7350-2, Big Sky, MT, March 2011

Wang, Q.; Noharet, B.; Junique, S. ; Agren, D. & Andersson, J. (2006). 1550 nm Transmissive/Reflective Surface-Normal Electroabsorption Modulator Arrays. *Electronics Letters*, Vol. 42, No. 1, January 2006, pp. 47-49, ISSN 0013-5194

Wang, Z.; Chang, S.; Ni, C. & Chen, Y. (2007). A High-Performance Ultracompact Optical Interleaver Based on Double-Ring Assisted Mach–Zehnder Interferometer. *IEEE Photonics Technology Letters,* Vol. 19, No. 14, July 2007, pp. 1072-1074, ISSN 1041-1135

Zhuang, L.; Roeloffzen, C.; Heideman R. ; Borreman A. ; Meijerink, A. & van Etten, W. (2007). Single-chip Ring Resonator-based 1x8 Optical Beamforming Network in CMOS-compatible Waveguide Technology. *IEEE Photonics Technology Letters,* Vol. 19, No. 13, July 2007, pp. 1130-1132, ISSN 1041-1135

Zhuang, L.; Roeloffzen, C.; Meijerink, A.; Burla, M.; Marpaung, D.; Leinse, A. ; Hoekman, M. ; Heideman, R. & van Etten, W. (2010). Novel Ring Resonator-Based Integrated Photonic Beamformer for Broadband Phased Array Receive Antennas— Part II: Experimental Prototype. *Journal of Lightwave Technology,* Vol. 28, No. 1, January 2010, pp. 19-31, ISSN 0733-8724

Part 4

Future Aeronautical Data Links

Future Aeronautical Communications: The Data Link Component

Nikos Fistas
EUROCONTROL[1]
E.U.

1. Introduction

This chapter presents an overview of the European activities in relation to the future aeronautical communications and in particular the new data link components. It presents the origins of the work, summarising the previous activities and it describes the three new data links that are being considered as well as the relevant activities undertaken in Europe. In addition, it briefly describes the airborne integration and the transition challenges that need to be investigated and resolved.

The content of this chapter is based on an article that was first published in the EUROCONTROL Skyway Magazine, No 54, in December 2010.

2. Background

The origin of the EUROCONTROL sponsored investigations concerning the future aeronautical communications can be traced back to ICAO's 11th Air Navigation Conference (AN-Conf/11) in 2003 (ICAO AN-Conf/11, 2003).

In its conclusions, AN-Conf/11 agreed that the aeronautical mobile communication infrastructure had to evolve in order to accommodate new functionalities and to provide the adequate capacity and quality of service required to support evolving air traffic management (ATM) requirements within the framework of the global ATM operational concept. Accordingly, the conference developed recommendations addressing the need for an evolutionary approach while ensuring the global interoperability of air/ground (a/g) communications, and requesting the investigation of the technology alternatives for future a/g communications and the standardisation of those selected.

The conference discussions stressed the requirement to maximise the use of systems already implemented, and highlighted the particular attention to be given to the careful utilisation of the (limited) available spectrum as well as to the appropriate consideration of transition aspects. In conclusion, the AN-Conf/ 11 emphasised the need for international cooperation, particularly in the field of air/ground communications.

In line with the conference recommendations, EUROCONTROL and the US Federal Aviation Administration (FAA) decided to establish a dedicated working arrangement

---

[1]The presented information expresses the views of the author and does not necessarily reflect EUROCONTROL's official policy.

(Action Plan 17 of the EUROCONTROL-FAA Memorandum of Cooperation) to carry out this work (EUROCONTROL/FAA Action Plan 17, 2007).

The Action Plan 17 (AP17) activities have been very closely coordinated with ICAO's Aeronautical Communications Panel (ACP) to facilitate global harmonisation. AP17 was a joint activity between FAA/NASA and EUROCONTROL. In Europe, France, Germany, Spain, Sweden and the UK have also been actively supporting and contributing to the European investigations. The AP17 work was concluded in 2007 and the outcome was used to plan the future activities in the context of the SESAR and NextGEN projects.

In Europe, the future communications infrastructure (FCI) work is now carried out in the context of the SESAR Programme, which will oversee the development of the required new generation of technological systems, components and operational procedures to support the future concepts as defined in the SESAR ATM Master Plan and Work Programme.

A key outcome of the AP17 activities was that there is no single technology which meets all expected future requirements across all operational flight domains. In addition to the importance of maximising the use of the existing infrastructure, the need to introduce technologies driven by clear operational requirements linked to tangible benefits, led to the conclusion that the FCI should be a "system of systems", integrating existing and new technological components. FCI should secure seamless continuation of operations supporting the current and future requirements, to safeguard investments in infrastructure and equipage, and to facilitate the required transitions.

In summary, the FCI work is built on the following assumptions:

- In the future operating concept, data becomes the primary mode of communications (voice will remain available for emergency communications).
- In the event of failure of data communications, voice is unlikely to be able to sustain operations at the same capacity level. Consequently, different data links may be needed in order to maintain capacity of operations.
- The future (2020+) system needs to support ATS (air traffic services) and AOC (airline operational communications) end-to-end communications, including ground/ground, air/ground and air/air.

The AP17 activities focused on air/ground data communications and analysed many candidate technologies. The technology investigations led to the following three proposals for new wireless data communications system developments (see Figure 1):

1. a ground-based, high-capacity, airport surface data link system, referred to as the aeronautical mobile airport communications system (AeroMACS);
2. a ground-based data link system for continental airspace in general, referred to as the L-band digital aeronautical communications system (LDACS);
3. a satellite-based data link system for the oceanic, remote (deserted) and continental environments (in the latter case complementing the terrestrial systems).

While for AeroMACS (the aeronautical mobile airport communications system), a specific existing standard has already been identified, for the other two technology developments, AP17 recommended further activities, to finalise the technical investigations and the selection and standardisation of the proposed systems.

It is important to note that a significant constraint in the technology investigations for aviation is the spectrum to be used. Communications supporting ATM exchanges fall into the category of "safety of life" communications, and as a result they have a special protection status in order to avoid interference.

However, this protection is applicable only to specific bands, and effectively there are three such bands: the VHF band, the L band and the C band. In Europe, the VHF band is a very

congested band, and there is no room for additional systems to operate. The VHF band was therefore not considered for use by any of the new systems, leaving the other two bands as the only candidates.
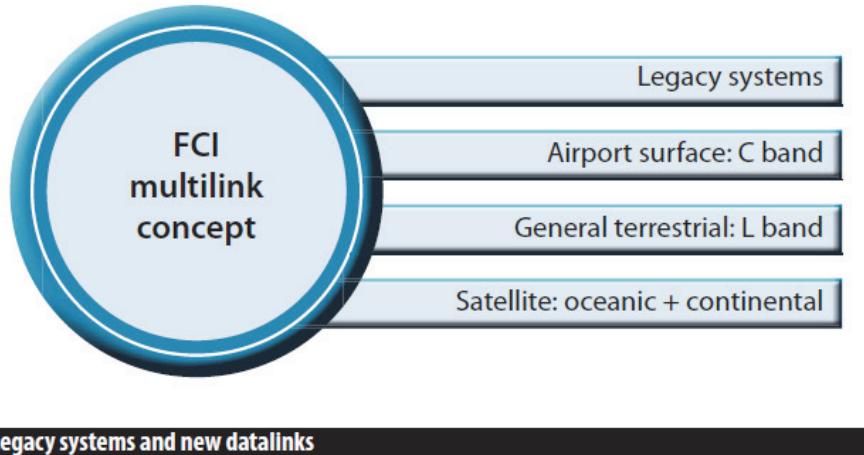

Fig. 1. FCI Multilink Concept.

3. Airport surface system: AeroMACS

AeroMACS is intended to support on-the-ground communication exchanges, particularly at busy airports. Whilst the aim is to support both ATS and AOC applications, it is expected that the AOC applications may be the driving element in the initial system implementations in Europe. In addition, especially in the US, the same standard is also being considered to support other aviation applications on the airport surface.

Considering the expected significant volumes of information to be carried and the short distances to be covered while the aircraft is on the ground, the aeronautical C band (5 GHz) has been selected, and an appropriate allocation for AM(R)S (aeronautical mobile route service allocation) was obtained by the International Telecommunication Union (ITU) in 2007. AeroMACS is based on the IEEE 802.16 WiMAX mobile communications standard, in order to benefit from commercial general telecom developments and minimise the required development resources.

In SESAR, there are two projects (P15.2.7 and P9.16) supporting the development of the AeroMACS system. These two projects aim to define and validate an international global aviation standard. The two projects will carry out analysis, simulations and testing, involving purpose-built system prototypes.

Project 15.2.7 addresses the overall system aspects and focuses on the ground system component, whilst project P9.16 focuses on the aircraft system aspects and investigates the airborne integration of the AeroMACS system.

In Europe, in addition to the SESAR projects, there is also the EU FP7 SANDRA project (SANDRA, 2009) that is actively pursuing the AeroMACS development. The SANDRA activities are closely coordinated with the SESAR projects activities aiming to avoid

duplication and maximise synergies. The AeroMACS development is also actively being pursued in the US supported by FAA/NASA.

Standardisation of AeroMACS is currently taking place in two closely cooperating groups in the European Organisation for Civil Aviation Equipment, EUROCAE, (WG82) (EUROCAE WG-82, 2010) and the US Radio Technical Commission for Aeronautics, RTCA, (SC223) (RTCA SC-223, 2010). These two groups are working on the development of the aviation-specific profile of the WIMAX standard to describe AeroMACS, and in the future the groups will also address the development of additional material (MOPS[2] and MASPS[3]).

In March 2011, the two groups converged on a joint proposal to WiMax Forum for the aviation profile. The features of this profile are now the basis for the development of prototypes in the context of SESAR and SANDRA activities in order to validate the profile and support the development of the required standards.

While EUROCAE and RTCA will cover the technical details of the proposed system, ICAO is also expected to be involved in the standardisation process, with high level documents such as SARPs[4]. A dedicated ICAO working group (WG S) has been formed under the Aeronautical Communications Panel (ACP) to develop the required ICAO documents, building upon the EUROCAE and RTCA relevant work.

4. Terrestrial air / ground system: LDACS

LDACS is a ground-based system using line-of-sight communications to support air/ground communication, in particular for en-route and TMA communications in continental airspace. The LDACS system is targeting operations in the L- band (1 GHz). This band is heavily utilised by navigation and surveillance aviation systems. However, considering the need to communicate over significant distances, the L-band was identified as the best compromise candidate band, primarily because of its acceptable propagation characteristics. A co-primary AM(R)S allocation was obtained from the ITU in 2007, which means that LDACS should not interfere with the other primary users of the band (navigation and surveillance systems).

The spectrum compatibility analysis is critical for LDACS and has to work in both directions: first not to hinder the operation of existing systems, but also to be able to operate in the presence of existing systems.

The AP17 investigations did not identify a commercially utilised system meeting all requirements, so they proposed to define a system based on features of some existing systems and reuse of previous developments. Following a trade-off analysis, two options for the LDACS were identified.

The first option (LDACS1) represents the state of the art in the commercial developments employing modern modulation techniques, and may lead to utilisation/adaptation of commercial products and standards. LDACS1 is based on a frequency division duplex (FDD) configuration utilizing OFDM modulation techniques, reservation based access control and advanced network protocols. This solution is a derivative of the B-AMC and TIA-902 (P34) technologies.

---

[2]Minimum Operational Performance Standards

[3]Minimum Aviation System Performance Specifications

[4]Standards And Recommended Practices

The second option (LDACS2) capitalises on experience from current aviation systems and standards such as VDL3[5], VDL4[6] and UAT[7]. LDACS2 is based on a time division duplex (TDD) configuration utilizing a binary modulation derivative of the implemented UAT system (CPFSK family) and existing commercial (e.g. GSM) systems and custom protocols for lower layers providing high quality-of-service management capability. This solution is a derivative of the LDL and AMACS technologies. The table below depicts the characteristics of the two options.

	Access scheme	Modulation type	Origins
LDACS1	FDD	OFDM	B-AMC, TIA 902 (P34)
LDACS2	TDD	CPFSK/GMSK type	LDL, AMACS

Table 1. LDACS (the L-band data link) candidates' key characteristics

The two LDACS options need further analysis, especially in terms of the spectral compatibility, before one of them is finally selected.

In SESAR, project P15.2.4 (Future Mobile Data Link system definition) is tasked with investigating the proposed LDACS options, selecting the most appropriate one, and developing the required standards. While some activities (Early Tasks) of the project are already ongoing, the full project activities will be starting in the second half of 2011.

The key tasks for the LDACS investigations are to define the interference environment and the criteria for the spectrum compatibility analysis. For these investigations, there will be development of prototypes and test beds to perform the required testing in a representative environment in order to validate the previous theoretical investigations and analysis carried out.

To ensure worldwide interoperability, the selected LDACS system will require global standards, which means that the decision for the LDACS system needs to be taken in the ICAO framework. Consequently, coordination with all other interested regions, such as the US, is very important. This is currently progressed under Action Plan 30, which has been established to continue the AP17 coordination. In the future, there will be a dedicated Coordination Plan addressing the cooperation between SESAR and NextGen under the EU /US Memorandum of Cooperation.

5. New satellite communication system

Satellite communications are very well placed to cover the large oceanic and remote airspaces. Currently, there are two ICAO standards for satellite communications, the INMARSAT3 and IRIDIUM systems. However, the performance requirements in the current ICAO satellite standards (AMS(R)S SARPs) are insufficient to cover the quality of service (QoS) requirements of the applications supporting the future operating concept. There is therefore a need to update the satellite SARPs to include more stringent performance requirements, to select a new technology and to develop the required standards to meet the updated requirements.

---

[5]VHF Digital Link 3

[6]VHF Digital Link 4

[7]Universal Access Transceiver

The need to select a new technology does not constitute an undesirable proliferation of technologies, as by the 2020+ timeframe all current aviation satellite systems will be reaching the end of their lifetime and new systems will have to be reconsidered in order to continue supporting the oceanic areas.

Furthermore, in order to support the new operating concepts, it is anticipated that multiple data links will be required to meet availability and other QoS requirements. It is therefore proposed that the satellite system should also be considered for use in continental airspace jointly with the terrestrial based systems to make it easier to meet the application requirements.

In Europe, the European Space Agency (ESA) has established the Iris project to facilitate the development of the future aviation satellite system. Iris is composed of technology studies such as ANTARES and THAUMAS investigating specific technical solutions. ANTARES is investigating the development of a new satellite communication system and THAUMAS is investigating the required evolution of the current INMARSAT Swiftbroadband system. In addition, Iris is also undertaking service provision studies investigating the suitable/likely model that is to be considered for the future satellite communication system.

The Iris programme is complemented by the SESAR P15.2.6 project (Future Satellite Communication system). The SESAR project focuses on the identification of the requirements, will support validation and verification activities in coordination with Iris and will undertake the required standardisation activities. In particular, the project will support the development of the required global ICAO standards in terms of SARPs and Technical Manuals.

The international aspect of satellite system standardisation is dealt with in a dedicated group, the NEXUS group, with voluntary contributions from all interested parties in Europe or outside (EUROCONTROL NEXUS). The first task of the group (technology independent) is to develop a coordinated proposal to ICAO for an updated version of the AMS(R)S SARPs with stringent performance requirements. In the future, the NEXUS group will continue with the development of a proposal for an ICAO Technical Manual describing a proposed system that will be meeting the performance requirements of the updated SARPs.

6. Airborne integration and transition considerations

As shown in Figure 2, if all the considered new data links will be implemented, then the future aircraft will have to support many and different communication modes. However it is already clear, that it will not be acceptable that the introduction of these new systems will imply the airborne integration of dedicated and standalone radio equipment and antennas. In addition to the weight and space limitations on the aircraft, the interference aspects make this option impossible to pursue.

Therefore the proposed new system developments need to be coupled with a new approach in order to address the airborne integration questions. The proposed Future Communications Infrastructure will only be feasible if a flexible airborne architecture is made possible. For such architecture, there are new promising technological developments that need to be considered. The use of software radios and the sharing of antennas or use of antennas able to operate in multiple frequency bands is a research field that needs further investigation.

In SESAR, project P9.44 (Flexible Communication Avionics) has two objectives. The first one is to investigate the technical and business feasibility and to identify solutions of new on-

board flexible radio architectures and equipment (such as, but not limited to, Software Defined Radio) which could support several or even all the future and legacy communication elements. The second objective, following an encouraging feasibility analysis will be to develop prototypes of candidate solutions for new on-board flexible radio equipment and to validate the appropriateness of this new technology as constituent of future certified avionics systems.
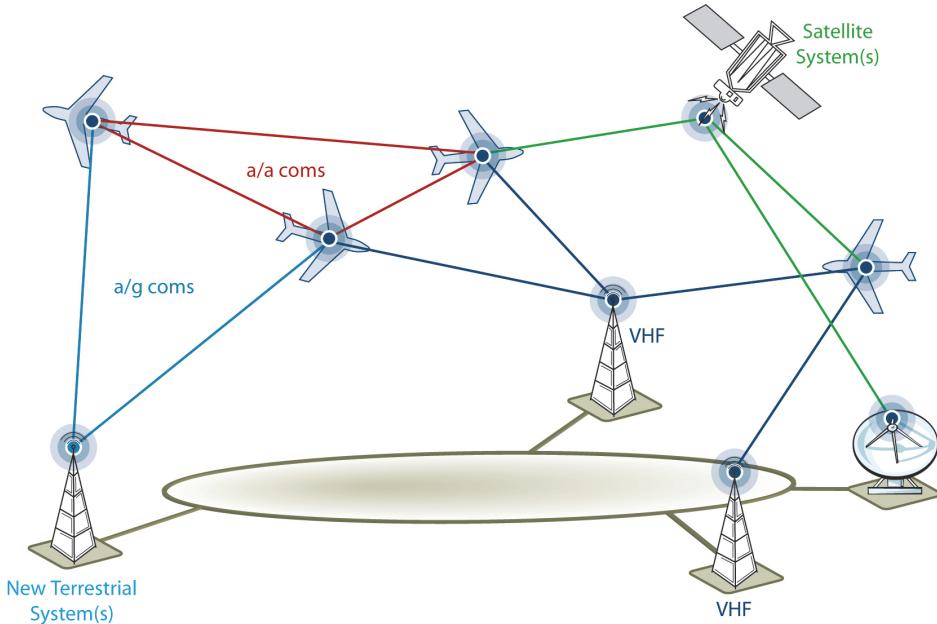

Fig. 2. FCI - a potential scenario for 2020+.

Another critical aspect that needs to be carefully addressed is the transition from the current to the future system. It is relatively easy to integrate the new technologies on new aircraft (forward fit) however retrofitting existing aircraft can be a challenging and sometimes an impossible task.

Therefore different scenarios need to be analysed in order to facilitate the implementation of the future systems. These scenarios need to ensure the optimisation of the use of the current (legacy) systems and only require the use of the new technologies for capabilities not being able to be supported by the legacy systems.

7. Conclusions

While significant work and achievements have already been accomplished, a lot of work remains to be done. Developing new technology solutions for aviation is a very lengthy and expensive process. There are many different factors which need to be properly aligned in order to bring about the successful implementation of a new system. Above all, any new system needs to be justified by new operational procedures and applications meeting future requirements.

New systems in aviation carry significant new costs for installation, and also additional costs for operation and maintenance, and they have to be offset by clear and measurable benefits. The proposed data links in the future communications infrastructure are not the drivers of the developments. They are the enablers of the required new concepts. However, in recognition of the long timescales required to introduce new systems in aviation, work needs to start many years in advance. The challenge in aviation is to select systems to be widely implemented at least a decade later, but which will remain capable of meeting evolving requirements.

In the technology development process, the international coordination is one of the prerequisites of success. It is critical to work with all interested parties from all regions. Timely availability of mature global standards will be critical in decision-taking and system implementation. EUROCONTROL is committed to working closely with ICAO and other interested parties to ensure that the appropriate systems are available to support the emerging concepts and operational requirements of the future 4D-trajectory-based operations.

8. References

EUROCAE WG-82. Working Group 82, Mobile Radio Communication Systems: Airport Surface Radio Link (WIMAX Aero), Info available at: www.eurocae.net/working-groups/wg-list/50-wg-82.html

EUROCONTROL/FAA Action Plan 17. (2007). Final Conclusions and Recommendations Report, Version 1.1, EUROCONTROL/FAA/NASA, November 2007, Available at: www.eurocontrol.int/communications/public/standard_page/General_FCI_Library.html

EUROCONTROL NEXUS. Information available at: www.eurocontrol.int/nexsat/public/standard_page/NEXUS.html

ICAO AN-Conf/11 (2003). Proceedings of Eleventh ICAO Air Navigation Conference, October 2003, Available at: www.icao.int

Iris. Satellite-based communication solution for the Single European Sky Air Traffic Management Research programme - Element 10 of the ESA ARTES programme, Info available at: www.telecom.esa.int/iris

NextGen. Next Generation Air Transportation System (NextGen), Info available at: www.faa.gov/nextgen

RTCA SC-223. Special Committee SC-223, Airport Surface Wireless Communications, Info available at: http://www.rtca.org/comm/Committee.cfm?id=133

SANDRA. Seamless Aeronautical Networking through integration of Data links Radios and Antennas (SANDRA), Info available at: www.sandra.aero

SESAR. Single European Sky ATM Research Programme (SESAR) Joint Undertaking, Info available at: www.sesarju.eu

Aeronautical Mobile Airport Communications System (AeroMACS)

James M. Budinger[1] and Edward Hall[2]
[1]NASA Glenn Research Center
[2]ITT Corporation
United States of America

1. Introduction

To help increase the capacity and efficiency of the nation's airports, a secure wideband wireless communications system is proposed for use on the airport surface. This chapter provides an overview of the research and development process for the Aeronautical Mobile Airport Communications System (AeroMACS). AeroMACS is based on a specific commercial profile of the Institute of Electrical and Electronics Engineers (IEEE) 802.16 standard known as Wireless Worldwide Interoperability for Microwave Access or WiMAX™. The chapter includes background on the need for global interoperability in air/ground data communications, describes potential AeroMACS applications, addresses allocated frequency spectrum constraints, summarizes the international standardization process, and provides findings and recommendations from the world's first AeroMACS prototype implemented in Cleveland, Ohio, USA.

1.1 Future communications for next generation air transportation

The highest concentration of sources, users, and stakeholders of information required for safe and regular flight operations occurs at the nation's airports. Of all flight domains within the national airspace system (NAS), the airport domain is the one where aircraft are in closest proximity to each other and to a wide variety of service and operational support vehicles, personnel, and infrastructure. Air traffic controllers, aircraft pilots, airline operators, ramp operators, aircraft service providers, and security, emergency, construction, snow removal, and deicing personnel all contribute to the safe and efficient operation of flights.

As the communications, navigation, and surveillance (CNS) facilities for air traffic management (ATM) at an airport grow in number and complexity, the need for communications network connectivity and data capacity increases. Over time, CNS infrastructure ages and requires more extensive and expensive monitoring, maintenance, repair or replacement. Airport construction and unexpected equipment outages also require temporary communications alternatives. Some typical examples of airport infrastructure, aircraft, service vehicles, and operators are shown in Figure 1.

Capacity growth in the nation's airports helps increase the total capacity of the NAS. But how can that growth occur while maintaining required safety, security, reliability, and diversity? A high-performance, cost-effective wireless communications network on the airport surface can provide part of the solution.

Fig. 1. Examples of typical airport infrastructure, aircraft, service vehicles, and operators that benefit from improved communications.

Through collaboration with the United States (U.S.), the Federal Aviation Administration (FAA) Headquarters in Washington, DC, the National Aeronautics and Space Administration (NASA) Glenn Research Center (Glenn) in Cleveland, Ohio, and its contractor, the ITT Corporation (ITT) in Fort Wayne, Indiana, are developing AeroMACS. AeroMACS is the first of three elements of the proposed future communications infrastructure (FCI)—a harmonization of future aeronautical air-to-ground (A/G) data communications capabilities intended to support the shared visions of the FAA's Next Generation (NextGen) Air Transportation System in the U.S. (FAA, 2011) and Europe's Single European Sky ATM Research (SESAR) program (SESAR, 2011).

AeroMACS offers the potential for transformational broadband secure wireless mobile data communications capabilities to future air traffic controllers, pilots, airlines, and airport operators on the airport surface. The unprecedented connectivity, bandwidth, and security afforded by AeroMACS have the potential to greatly enhance the safety and regularity of flight operations in the future.

2. Call for global harmonization

This section describes the steps that led to joint recommendations between the U.S. and Europe for a future wireless communications network on the airport surface. In the early 2000s, the International Civil Aviation Organization (ICAO) Aeronautical Communications Panel (ACP) recognized that the very high frequency (VHF) band allocated globally for A/G voice and data communications for ATM was beginning to reach saturation. The problem was characterized at the time as being more severe in Europe than in the U.S. However, both had taken steps to significantly reduce VHF channel spacing (from 50 kHz to 25 kHz in the U.S. and from 25 kHz to 8.33 kHz in Europe). This reduction allows more simultaneous voice and data services in the crowded VHF spectrum. Various proposals for digital A/G datalinks from individual countries obtained ICAO approval independently. But none achieved global endorsement.

The call to action came from ICAO's Eleventh Air Navigation Conference (ANC-11) held in Montréal, Quebec, Canada in late 2003. ANC-11 advanced the operational concept of global

ATM as guidance for the development of future ATM-related service provisions through the year 2025 and beyond. The official report of the "Technical and Operational Matters in Air Traffic Control Committee" included several observations regarding the state of global aviation communications (ICAO ANC, 2003). Those included the need for the aeronautical mobile communications infrastructure to evolve in order to accommodate new functions, the gradual introduction of data communications to complement and eventually to replace voice for routine communications, and the universally recognized benefits of harmonization and global interoperability of A/G communications.

The committee made specific recommendations to develop an evolutionary approach for global interoperability of A/G communications (Recommendation 7/3), conduct an investigation of future technology alternatives (Recommendation 7/4), and prove compliance with certain minimum criteria before undertaking future standardization of aeronautical communications systems (Recommendation 7/5).

2.1 Future communications study

These recommendations helped establish the goals for the Future Communications Study (FCS), a joint investigation by the FAA and EUROCONTROL also referred to as Action Plan 17 (AP-17) under the Memorandum of Cooperation between the two organizations. AP-17 was approved in early 2004 and modified over the course of the 4-year FCS (Fistas et al., 2007). Under the FCS, the Communications Operating Concepts and Requirements (COCR) for future A/G data communications was developed jointly and revised once (version 2.0) (ICAO COCR, 2007). The COCR provides the shared vision for future ATM concepts of operations and services in all flight domains. Two phases of development were considered.

The first phase (roughly 2005 through 2020) envisions increased use of data communications, but within the current tactical aircraft management practices at the time. The concepts identified in the second phase (roughly 2020 and beyond) anticipate a paradigm shift towards ATM based primarily on data communications in all flight domains, with use of voice intervention by exception. The COCR served as the basis for selecting technologies for the future radio system (FRS), the A/G portion of the overall FCI. The FCS technology assessment directly implemented a key recommendation of ANC-11.

"Recommendation 7/4 — Investigation of future technology alternatives for air-ground communications: That ICAO

a.　investigate new terrestrial and satellite-based technologies, on the basis of their potential for ICAO standardization for aeronautical mobile communications use, taking into account the safety-critical standards of aviation and the associated cost issues;

b.　continue evolutionary development of existing standardized ICAO technologies with a view to increasing their efficiency and performance; and

c.　assess the needs for additional aeronautical spectrum to meet requirements for increased communications capacity and new applications, and assist States in securing appropriate additional allocations by the ITU (ICAO ANC, 2003)."

2.2 Common recommendations for future communications infrastructure

Beginning in 2004, NASA Glenn and its contractor, ITT, conducted the FCS technology assessment for the U.S., evaluating the technical and operational capabilities of over 60 different commercial, public safety, and Government communications services and

standards for applicability to the COCR. The U.S. technology assessment was conducted in close cooperation with EUROCONTROL and their contractor, QinetiQ. The process was conducted in multiple phases. The first of these, technology pre-screening, provided an initial down-selection against detailed functional and performance evaluation criteria. The second phase included detailed investigations of a smaller set of candidates. Simulation and evaluation in the third phase led to a harmonized shortlist of common recommendations. The process is illustrated in Figure 2 (Gilbert et al., 2008).
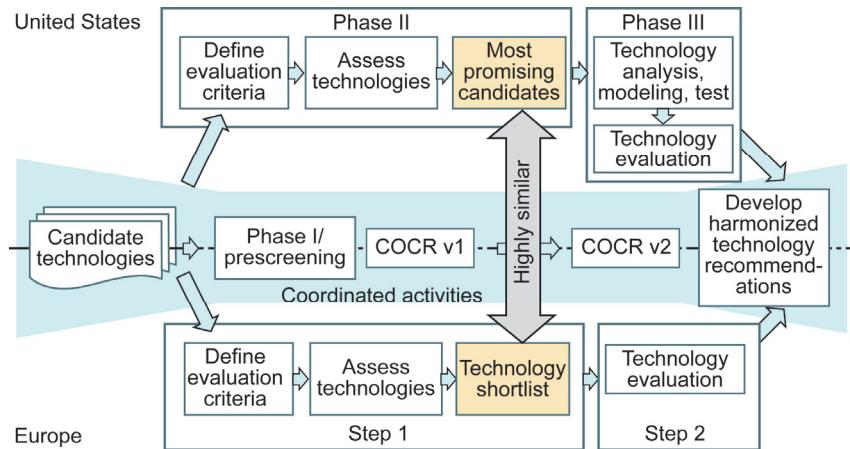

Fig. 2. The technology assessment process used in the Future Communications Study.

The international harmonization process was carried out over multiple meetings of ICAO's Aeronautical Communications Panel (ACP) Communications Working Groups (WGC-8 through WGC-11) and Working Group Technology (WGT) to establish common solutions for future A/G data communications in the 2020 timeframe (ICAO WGC, 2006) (Phillips et al., 2007). An underlying objective of the FCS technology assessment was to maximize existing technologies and standards and minimize any modifications to each. This approach leverages existing commercial industry resources invested in developing and standardizing the technology and can expedite ICAO approval as an international aviation standard.

The FCS technology assessment considered technology candidates as elements of FCI in three flight domains—continental (i.e., enroute airspace within line of sight of terrestrial air traffic control (ATC) communications facilities), oceanic and remote airspaces (i.e., enroute airspace beyond line of sight of terrestrial facilities), and airport (i.e., pre-departure and post-arrival on the surface). The common shortlist of technologies recommended for further evaluation through prototype developments was approved by ACP in April 2008 at the second Working Group of the Whole (WGW-2) and is summarized in Figure 3 (ICAO WGW, 2008). Gilbert et al., 2006 provides details regarding evaluation of IEEE 802.16e for the airport surface.

The common recommendation to be used as the starting point for aeronautical wireless mobile data communications on the airport surface was the 2005 version of the IEEE standard for local and metropolitan area networks, IEEE 802.16e. AeroMACS, the first element of the FCI, is based on the most current version of this standard, IEEE 802.16-2009, Part 16: Air Interface for Broadband Wireless Access Systems (IEEE, 2009).
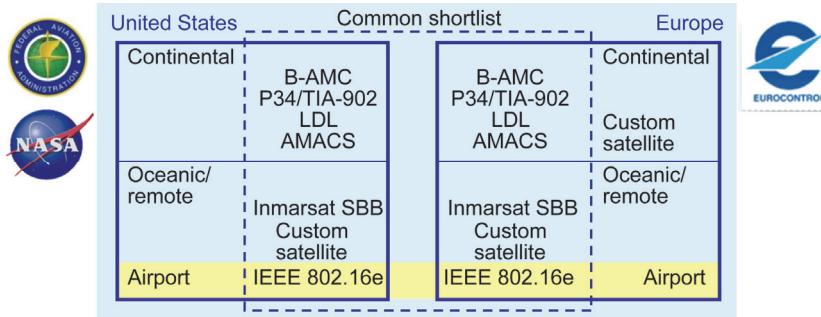
Fig. 3. The common technology recommendations of the Future Communications Study.

Throughout this chapter, the term "IEEE 802.16" will refer to the 2009 version of that standard. The evolving standard is well suited for implementation below 11 GHz. The amendment for mobility uses 512 subcarrier (in 5-MHz channel) orthogonal frequency division multiple access (OFDMA) modulation and supports multiple channel bandwidths from 1.25- to 20-MHz, with peak duplex data rates above 50 Mbps. Table 1 highlights some features of the IEEE 802.16 mobile standard that makes it attractive for use on the airport surface.

A specific WiMAX Forum® profile of the IEEE 802.16 standard is proposed for AeroMACS. This enables the aviation community to leverage extensive international standards collaboration and commercially provided components and services (WiMAX Forum®, 2011a). Section 5 provides more details regarding the WiMAX™ profile selected for AeroMACS.

Feature	Advantages
Mobility	Supports vehicle speeds of up to 120 km/hr , sufficient for aircraft taxiing and emergency surface vehicle speeds
Range	Covers up to ~10 km in line-of-sight (LOS) communications, sufficient to cover most airports
Link Obstruction Tolerance	Exploits multipath to enable non line-of-site (NLOS) communications
Quality of Service (QoS)	Enables QoS based on throughput rate, packet error rate deletion, scheduling, time delay and jitter, resource management
Scalability	Includes flexible bandwidth and channelization options to enables network growth on demand
Security	Includes mechanisms for authentication, authorization, strong encryption, digital certificates, and fast handovers
Privacy	Supports private Virtual Local Area Networks (VLANs)
Open Sourced	Leverages modern communications technologies and supports modern Internet-based network protocols
Cost Efficiency	Via commercial standards and components, industry capabilities, and reduced physical infrastructure compared with buried cable

Table 1. Features of IEEE 802.16 desirable for implementation of AeroMACS networks

3. Potential AeroMACS configuration and applications

An AeroMACS based on the WiMAX™ standard for local area networks can potentially support a wide variety of voice, video, and data communications and information exchanges among mobile users at the airport. The airport CNS infrastructure that supports ATM and ATC on the airport surface can also benefit from secure wireless communications by improving availability and diversity.

A wideband communications network can enable sharing of graphical data and near real-time video to significantly increase situational awareness, improve surface traffic movement to reduce congestion and delays, and help prevent runway incursions. AeroMACS can provide temporary communications capabilities during construction or outages, and can reduce the cost of connectivity in comparison to underground cabling. A broadband wireless communications system like AeroMACS can enhance collaborative decision making, ease updating of large databases and loading of flight plans into flight management system (FMS) avionics, and enable aircraft access to system wide information management (SWIM) services for delivery of time-critical advisory information to the cockpit.

3.1 Proposed AeroMACS network configuration

To provide services to a potentially large number of mobile users and fixed assets, a standard WiMAX™ network architecture is proposed for AeroMACS. One or more base stations are required to provide required coverage, availability, and security. Figure 4 illustrates a notional AeroMACS network deployed at an airport.
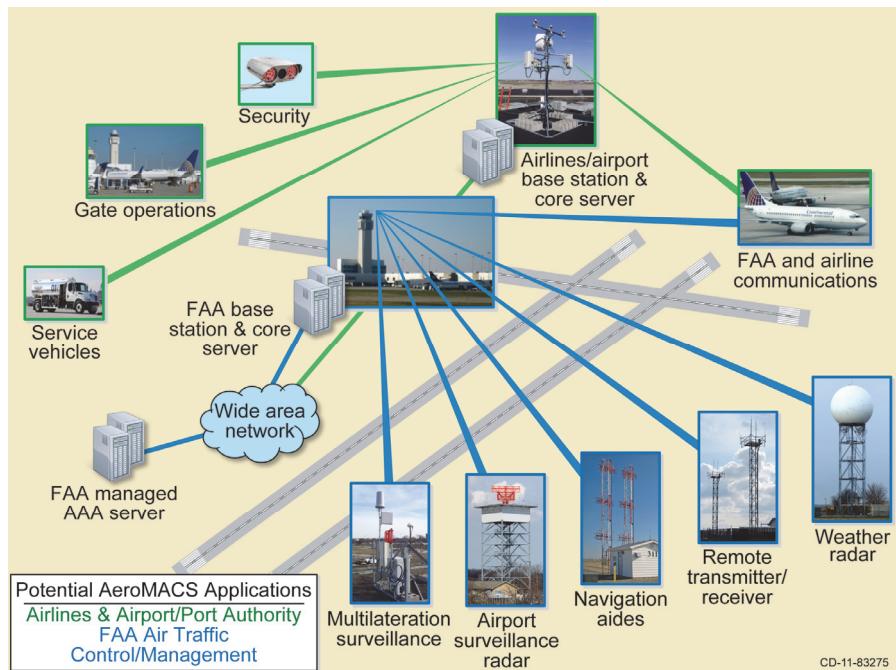

Fig. 4. Notional AeroMACS network configuration and potential applications.

In this notional network configuration, air traffic control and management services can be physically isolated from airlines and airport/port authority services if required. However, WiMAX™ networks have the capability to integrate multiple services while preserving the desired security and quality of service provisions of each.

3.2 Categories of potential AeroMACS services

The potential services and applications provided by AeroMACS can be grouped into three major categories: ATC/ATM and infrastructure, airline operations, and airport and/or port authority operations (Budinger et al., 2010). Within these broad categories, the data communications services and applications can be described as either fixed or mobile, based on the mobility of the end user. However, because of operational constraints on the international frequency spectrum allocated for AeroMACS (described in section 4), only those services that can directly impact the safety and regularity of flight are candidates for provision by AeroMACS. Some examples of potential AeroMACS services and applications are listed in Table 2.

FAA Air Traffic Control and Infrastructure Applications Examples	
• Selected air traffic control (ATC) and air traffic management (ATM)	Mobile
• Surface communications, navigation, surveillance (CNS), weather sensors	Fixed
Passenger and Cargo Airline Applications Examples	
• Aeronautical operational control (AOC)	Mobile
• Advisory information	Mobile
• Aeronautical information services (AIS)	
• Meteorological (MET) data services	
• System wide information management (SWIM)	
• Airline administrative communications (AAC)	Mobile
Airport Operator/Port Authority Applications Examples	
• Security video	Fixed
• Routine and emergency operations	Mobile
• Aircraft de-icing and snow removal	Mobile

Table 2. Examples of potential AeroMACS services and applications.

3.2.1 Potential air traffic applications

Many candidate mobile ATC/ATM applications are under consideration for future provision via AeroMACS (Apaza, 2010). These include selected messages that are currently conveyed over the aircraft communications addressing and reporting system (ACARS) (e.g., pre-departure clearance (PDC)), selected controller pilot data link communications (CPDLC) messages (e.g., 4-dimensional trajectory negotiations (4D-TRAD)), selected COCR services (e.g., surface information guidance (D-SIG)), and other safety-critical applications (e.g., activate runway lighting systems from the cockpit (D-LIGHTING)). Potential fixed infrastructure applications in the U.S. include communications (e.g., controller-to-pilot voice via remote transmit receiver (RTR)), navigation aids (e.g., instrument landing system data for glide slope and visibility data for runway visual range), and surveillance (e.g., airport surface movement detection and airport surveillance radar (ASR)). AeroMACS can also be used to convey electronic equipment performance data for remote maintenance and

monitoring (RMM). Most of these existing applications are fixed point-to-point and use voice grade circuits. AeroMACS offers a flexible alternative to guided media (e.g., copper and fiber optic cable). However, the FAA may require separation of these services from the airline and airport services, which are described in the next two subsections.

3.2.2 Potential airline and advisory applications

Mobile AIS/MET services have the potential to become significant drivers of AeroMACS design because of several high-volume data base synchronization services that would benefit from AeroMACS implementation (Apaza, 2010). These include the AIS baseline synchronization service (e.g., uploading flight plans to the FMS and updating terrain and global positioning satellite (GPS) navigational databases and aerodrome charts to electronic flight bag (EFB)), data delivery to the cockpit (e.g. data link aeronautical update services (D-AUS), and airport/runway configuration information (D-OTIS)), and convective weather information (e.g., graphical forecast meteorological information and graphical turbulence guidance (GTG) data and maps).

Passenger and cargo airlines provide another significant source of data and voice applications for potential integration over AeroMACS. These include ground operations and services (e.g., coordination of refueling and deicing operations), sharing of maintenance information (e.g., offload of flight operational quality assurance (FOQA) data), and aircraft and company operations (e.g., updates to flight operations manuals and weight and balance information required for takeoff).

3.2.3 Potential airport operator applications

The airport or port authority operations provide the final category of potential applications for AeroMACS (Apaza, 2010). These are dominated by video applications required for safety services (e.g., fixed surveillance cameras and in-vehicle and portable mobile cameras for live video feeds and voice communications with central control during snow removal, de-icing, security, fire and rescue operations). Finally, AeroMACS can also help ensure compliance with regulations for safety self-inspection (e.g., reporting status of airport runway and taxiway lights and monitoring and maintenance of navigational aids and time critical airfield signage). The full range of candidate applications and services for AeroMACS is under investigation in both the U.S. and Europe (Wargo & Apaza, 2011).

Many of these services and applications are currently provided to mobile users through a mix of VHF voice and data links, land mobile radio services, and commercial local area wireless networks. The fixed communications services and applications at airports are typically implemented via buried copper and fiber optic cables. AeroMACS offers the potential for integration of multiple services into a common broadband wireless network that also securely isolates the applications from each other.

The first safety-critical application expected to migrate to AeroMACS in the U.S. is airport surface detection equipment model X (ASDE-X). For ASDE-X, AeroMACS provides wireless interconnection of multilateration (MLAT) sensors distributed across the airport surface. MLAT data is combined with surface movement radar data and aircraft transponder information to display detailed information about aircraft position (Sensis, 2011).

The deployment of AeroMACS infrastructure at an airport to enable the migration or augmentation of one of more existing services opens the potential for many additional services, especially those that require wider bandwidth, such as graphical information delivery and video services.

4. Spectrum considerations

This section describes the process leading to an international frequency spectrum allocation for AeroMACS, and modeling to ensure compatibility with other co-allocations in the band.

4.1 Channel modeling

The provision of a new international frequency spectrum allocation for the future airport surface wireless data communications system was supported by C-band channel modeling and service bandwidth estimation studies. Signal propagation research and channel sounding measurements at 5091- to 5150-MHz were performed by Ohio University and NASA Glenn at airports in the U.S. (Matolak, 2007). Measurements were taken at representative large, medium and small (general aviation) airports.

Thousands of power delay profiles (PDPs) were taken at each airport, along with received signal strength (RSS). In general, wireless communications networks at large airports will experience the most areas of multipath fading and non-line-of-sight (NLOS) conditions. Figure 5 illustrates an example of the time evolution of an NLOS PDP, taken from measurements at JFK Airport.
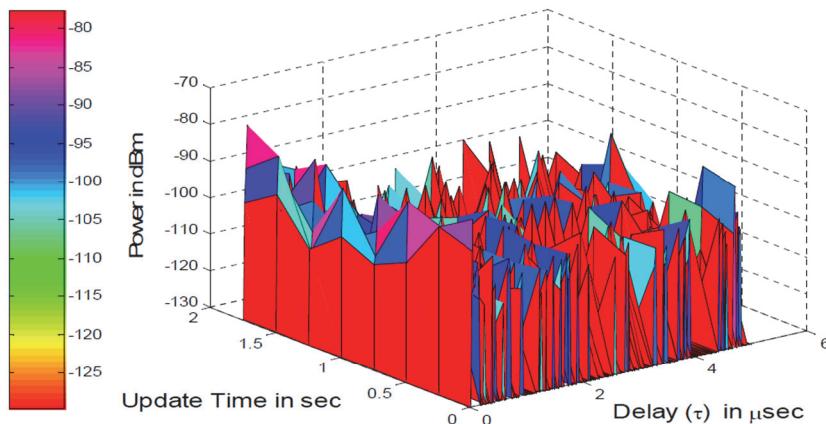

Fig. 5. An example of a power delay profile versus time.

The example shows how the received components fade in time. Fades of more than 10 dB are evident. The PDP and receive signal strength indication (RSSI) measurements enabled characterization of propagation path loss, fading channel amplitude statistics, multipath persistence and channel statistical non-stationarities, and fading rate. Observations during measurements also revealed highly non-isotropic scattering. The study concluded that the airport surface channel is very dispersive for bandwidths above about 1 MHz and that fading is very dynamic and in some cases severe.

These characteristics were used to develop statistically nonstationary tapped delay line channel models for both high fidelity (HF) and sufficient fidelity (SF). Because of the complexity of the HF models, the study recommended that the SF models be used to evaluate the performance of the proposed IEEE 802.16 systems in the airport surface environment.

4.2 Bandwidth estimation for proposed spectrum allocation

Studies to estimate the bandwidth required to provide the potential AeroMACS applications such as those identified in section 3 were conducted in collaboration with the FAA by both NASA and the MITRE Corporation Center for Advanced Aviation System Development (CAASD). An early NASA/FAA study estimated the FAA's existing and anticipated data requirements for instrument landing systems, radar systems, runway visual range, visual aids, and A/G communications (Apaza, 2004). The highest requirements for wireless communications from airlines and port authorities included communications with ground maintenance crews and airport security.

A later study conducted by NASA Glenn estimated additional bandwidth requirements to accommodate wake vortex sensing (to potentially enable closer spacing between arriving aircraft), and the overhead associated with security provisioning features of the IEEE 802.16 standard (Kerczewski, 2006).

In a series of studies conducted for the FAA from 2004 to 2008, MITRE CAASD established and refined estimates of the aggregate data rate requirements for a high-data-rate surface wireless network called airport network and location equipment (ANLE) (Gheorghisor, 2008). In alignment with the COCR, these studies addressed potential requirements through 2020 (Phase 1) and beyond 2020 (Phase 2). The bandwidth requirements for proposed mobile and fixed applications using an IEEE-802.16-based system were estimated for both low-density and high-density airports.

The highest total aggregate data capacity requirements for fixed and mobile applications is based on large airports (e.g., Dallas Ft. Worth (DFW)) with a terminal radar approach control (TRACON) ATC facility not collocated with an ATC tower (ATCT). ANLE was envisioned primarily to provide mobile communications with aircraft, but also to support classes of sensors and other fixed and mobile applications within the same network.

4.2.1 Aggregate data rate for mobile applications

Aggregate data requirements for ANLE were estimated for the following categories of mobile applications for the Phase 2 timeframe, listed in decreasing magnitude:

- Large file transfers from AOC to onboard electronic flight bags (EFBs) such as database updates and graphical weather
- Monitoring and controlling the physical security of aircraft including the provision of real-time video transmission from the cockpit
- Integration and dissemination of situational awareness information to moving aircraft and other vehicles
- Voice over Internet protocol (VoIP) among airline and airport personnel
- Radio frequency identification (RFID) for luggage and other assets.

The estimated aggregate data rate requirement for these mobile applications is nearly 20 Mbps. AOC data accounts for more than half of that.

4.2.2 Aggregate data rate for fixed applications

Estimates for the following categories of fixed applications for the Phase 2 timeframe, listed in decreasing magnitude are

- Communications from sensors for video surveillance and navigational aids to the TRACON
- TRACON-to-ATCT video, voice, and data communications

- Diversity path for ATC voice to the RTR
- Distribution of weather data products
- Surveillance data from surface radars and ASDE-X sensors.

The estimated aggregate data rate requirement for these fixed applications is over 52 Mbps. The combination of video surveillance and sensors and TRACON-to-ATCT data communications account for about 80% of the total.

The combined mobile and fixed data requirements provided the basis for estimating the total amount of radio spectrum needed for the operation of ANLE, now referred to AeroMACS. Based on analysis of an IEEE 802.16 system, two different base station channel bandwidth configurations (multiple 10-MHz and 20-MHz channels) and modulation techniques, an upper bound of 60 MHz of new spectrum was estimated in order to support the envisioned applications in the 2020 timeframe and beyond. The ITU-R expects that 60 to 100 MHz of spectrum will be required for the future surface domain (ITU-R, 2007).

4.3 International spectrum allocation

At the International Telecommunications Union World Radiocommunication Conference held in late 2007 (WRC-07), Agenda Item 1.6 invited participants "to consider allocations for the aeronautical mobile route service (AM(R)S) in parts of the bands between 108 MHz to 6 GHz, and to study current frequency allocations that will support the modernization of civil aviation telecommunication systems." At the conclusion of WRC-07, a new AM(R)S co-primary allocation in the 5091-5150 MHz band was added to the International Table of Frequency Allocations. The new allocation is limited to surface applications at airports. This allocation is in a region of the frequency spectrum commonly referred to as C-band.

This specific 59 MHz of spectrum is also referred to as the microwave landing system (MLS) extension band. MLS carries an aeronautical radio navigation services (ARNS) allocation. The WRC-07 decision on Agenda Item 1.6 essentially removed the prior limitation for support of ARNS only. Along with the existing MLS and new AeroMACS services, the other co-primary service allocations in this band include Earth-to-Space satellite feeder links for non-geostationary orbiting (GSO) mobile satellite service (MSS), and new co-allocations for aeronautical mobile telemetry (AMT) used with research aircraft during test flights and an aeronautical mobile service (AMS) limited to aeronautical security (AS).

The AM(R)S communications are defined as safety communications requiring high integrity and rapid response. Generally these include ATC and those AOC communications that support safety and regularity of flight (Biggs, 2008). In the U.S., AeroMACS networks are expected to be approved for both mobile and fixed applications that directly support safety and regularity of flight. AeroMACS services can be provided to aircraft anywhere on the airport surface, as long as wheels are in contact with the surface. AeroMACS can also be used for communications with a variety of service vehicles and airport infrastructure that directly support safety and regularity of flight.

The protected allocation for AM(R)S in this portion of C-band enables ICAO to approve international standards for AeroMACS wireless mobile communications networks on the airport surface. Based on expectation of high demand for AeroMACS services, Agenda Item 1.4 for WRC-12 will consider additional allocation of AM(R)S spectrum within the 5000-5030 MHz band.

4.4 Modeling for interference compliance

The co-allocation for AeroMACS at WRC-07 includes provisions to limit interference with other co-primary terrestrial services—MLS, AMT, and MSS feeder links. In the U.S.,

essentially no airports use the MLS for precision landing assistance. That need has been largely met through the wide area augmentation system (WAAS) that is based on GPS data. A limited number of airports in Europe use MLS. At those airports, coordination for equitable sharing of the 59-MHz allocation will be required to prevent mutual interference. In similar fashion, civilian airports near the specific locations where AMT is used on test aircraft will need to coordinate on the use of specific AeroMACS channels and AMT transmissions in order to limit potential interference. However, potential interference from hundreds of AeroMACS-equipped airports across the continents into MSS feeder link receivers on orbiting satellites is global in nature. In specific, the potential for co-channel interference from AeroMACS into the Globalstar MSS feeder link receivers must be mitigated through practical limits, international standards, and compliant implementations across the nations' airports.

NASA Glenn is modeling the interference caused by AeroMACS in order to help establish practical limits on the total instantaneous power that could eventually be radiated from hundreds of airports across the NAS (Wilson & Kerczewski, 2011). In order to ensure that the MSS feeder link threshold is not exceeded, the total radiated power recommended for each potential AeroMACS-equipped airport must take into consideration the total radiated power from all potential AeroMACS-equipped airports across the NAS. NASA Glenn uses Visualyse Professional Version 7 software from Transfinite Systems Limited to model the potential interference.

Figure 6 illustrates the aggregate interference power at a single Globalstar satellite receiver orbiting at 1414-km from AeroMACS emissions at a total of 757 towered airports across the U.S. and the Caribbean, including 34 in Canada, and 20 in Mexico. For this condition, the model assumes each airport radiates 5.8-W omni-directionally in the 20-MHz channel that spans the Global receiver's 1.23 MHz bandwidth.

|	
	−157.3
	−159.3
	−161.3
	−163.3
	−165.3
	−172.4
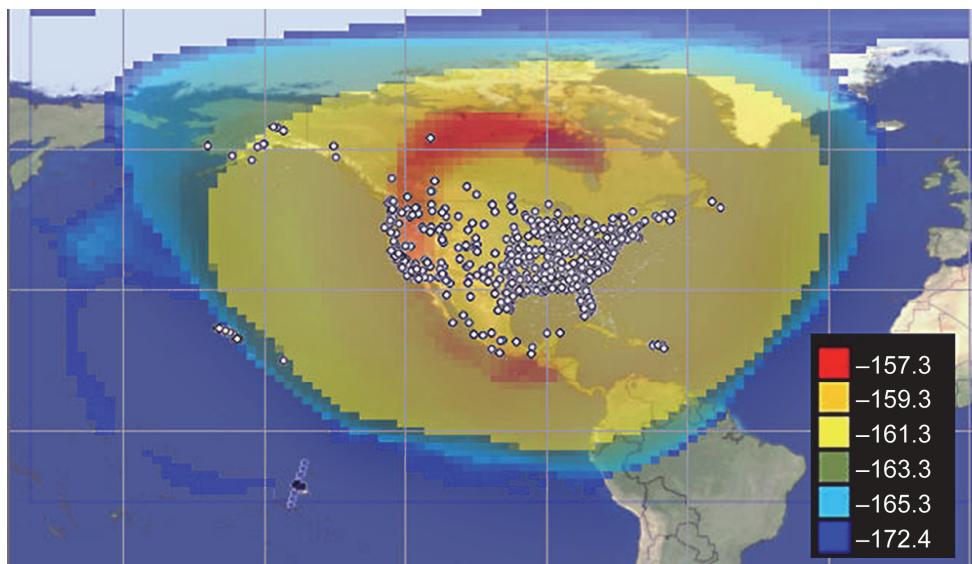
Fig. 6. Modeled interference power distribution from 757 AeroMACS-equipped airports in North America as seen at a Globalstar receiver orbiting 1414 m above the Earth's surface.

Based on an interpretation of the final resolutions of WRC-07, the AM(R)S co-allocated services must not increase the thermal noise temperature of Globalstar feeder link receivers by more than 2% (Gheorghisor et al., 2009). This corresponds to a threshold of -157.3 dBW for total interference power from AeroMACS into Globalstar feeder link receivers. In order to prevent the interference power from exceeding this threshold at any point in the Globalstar receiver orbit, the model shows that the omni-directional transmitters at each of the 757 airports needs to be limited to 799-, 401-, and 201-mW for 20-, 10-, and 5-MHz channels, respectively (Wilson & Kerczewski, 2011).

Further enhancements to the realism of the airport infrastructure modeling and correlation with experimental performance measurement data from the AeroMACS prototype are underway at NASA Glenn. Enhanced models will be used to develop a final set of recommendations on AeroMACS radiated power limits based on 5-MHz channels (Gheorghisor et al. 2011).

4.5 Proposed AeroMACS channelization

The location of AeroMACS channels within the 5091- to 5150-MHz allocation takes into consideration a number of factors. Among those are efficient utilization of current and potential future spectrum allocations; guard bands to limit out-of-band radiated power; anticipated number of AeroMACS BSs and SSs; practical limits on frequency spectrum reuse; the bandwidth requirements of potential AeroMACS applications described previously; and compliance with WiMAX Forum® standards.

The channel plan illustrated in Figure 7 shows the recommended AeroMACS channel plan. It includes 5-MHz channels on equally spaced center frequencies from 5095-MHz to 5145-MHz. Assuming coordination with other aviation allocations in the band directly below 5091 MHz (to limit the effects of interference) enables up to 11 separate AeroMACS channels. This plan can be extended to accommodate additional 5-MHz channels for a future allocation within the 5000-5030 MHz spectrum (Budinger et al., 2010).
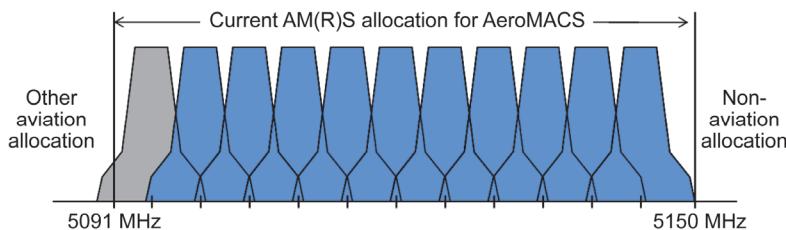

Fig. 7. Proposed AeroMACS channel plan for 5091-5150 MHz allocation.

5. International standards process

This section summarizes the process, findings, and recommendations provided by the FAA, NASA Glenn and ITT to advance the application of a specific IEEE 802.16 profile as the basis for the AeroMACS standard. A standard profile for AeroMACS ensures that all stakeholders—test equipment vendors, integrated circuit vendors, as well as the aviation industry—are capable of supporting AeroMACS development, and that global deployments will be interoperable. The profile is used as a guide for development of minimum operational performance standards (MOPS) for AeroMACS avionics.

In the U.S., an RTCA Special Committee on Airport Surface Wireless Communications, SC-223, was established in July 2009 to develop the AeroMACS profile and MOPS (RTCA SC-223, 2011). The U.S. final draft profile was completed at the end of 2010 and the MOPS document is scheduled to complete by the end of 2011. The AeroMACS profile and MOPS are developed in close coordination with EUROCAE Working Group WG-82 in Europe. Common AeroMACS standards in the U.S. and Europe are requested by ICAO in part to be responsive to the recommendation of ANC-11 for global interoperability and to help expedite ICAO approval of international AeroMACS standards.

The AeroMACS profile closely follows the format and substance of profiles developed by the WiMAX Forum® for commercial and industrial use. The WiMAX Forum® is an industry consortium whose primary technical function is to develop the technical specifications underlying WiMAX Forum Certified™ products. An ad-hoc joint committee was established between RTCA SC-223 and the WiMAX Forum® in August, 2010, to facilitate development of an AeroMACS profile. The profile is expected to be incorporated as one of several WiMAX Forum Certified™ profiles.

5.1 WiMAX forum® profiles

The initial RTCA AeroMACS profile is based on the WiMAX Forum® Mobile System Profile Specification Release 1.0 because it is currently the only release recommended by the WiMAX Forum® for hardware certification use. Release 1.5 has been approved by WiMAX Forum® but is not implemented for hardware certification because the IEEE 802.16m amendment is expected to be implemented soon via profile Release 2.0. The RTCA SC-223 and EUROCAE WG-82 decided jointly not to implement features of the upcoming profile Release 2.0 at this time. Thus, the AeroMACS standard is currently based on Release 1.0.

Release 1.0 is published in three main parts: (1) COMMON Part, (2) Time Division Duplex (TDD) Part, and (3) Frequency Division Duplex (FDD). However, AeroMACS is recommended to be a TDD-only system, so only the first two parts of the WiMAX Forum® profile are applied to AeroMACS (WiMAX Forum®, 2011a).

5.2 Joint RTCA – EUROCAE process

The AeroMACS profile has been developed through a series of RTCA and EUROCAE meetings and telephone conferences, often with WiMAX Forum® participation. SC-223 and WG-82 leadership participate in most plenary meetings of each other's organizations.

A joint RTCA and EUROCAE meeting was held in Brussels, Belgium in late September, 2010 with participation by members of the WiMAX Forum® via telephone conference in which many profile parameter settings were established for AeroMACS. A fully harmonized profile was established during the RTCA SC-223 Plenary Meeting #8 in November, 2010. This harmonized profile is available on the RTCA SC-223 website; however, permission from RTCA is required to access the workspace where these documents are posted. The profile description at this site includes a rationale statement for each chosen setting.

The joint AeroMACS profile completed in December 2010 is considered as the RTCA "final draft" version. EUROCAE plans to continue their studies throughout 2011, leading to a "final joint profile" by the end of 2011. The final joint profile may differ from the 2010 final draft profile based on results of the EUROCAE studies. EUROCAE plans to complete validation tests before publishing a final AeroMACS profile by the end of 2013.

Commercial WiMAX™ networks have been successfully deployed in 140 countries as of May 2009. This global acceptance of the WiMAX™ standard, and U.S. interoperability with

the AeroMACS standard approved by EUROCAE in Europe, is expected to ease acceptance and ICAO approval of the global AeroMACS standard for aeronautical mobile applications on the airport surface.

6. AeroMACS prototype network

This final section of the chapter discusses the development and evaluation of an AeroMACS prototype network. The FAA-sponsored AeroMACS research is identified in the FAA's NextGen Implementation Plan for 2009 and 2010. A reimbursable Space Act Agreement between NASA Glenn and the FAA enables collaboration between these two agencies and contracted support from ITT for AeroMACS research, development, and service demonstrations.

The world's first AeroMACS prototype was completed in late 2009 for validation of airport surface concepts and verification of communications performance requirements. The AeroMACS prototype is deployed within the Communications, Navigation, and Surveillance (CNS) Test Bed located at NASA Glenn and adjacent Cleveland Hopkins International Airport (CLE). In the following subsections, a description is provided of the AeroMACS prototype and some of the practical technical tradeoffs associated with coverage, cost, and performance, followed by initial results of AeroMACS performance experiments.

Full details of the multiple-year AeroMACS research, development, and experimental effort are available in a two-volume final report (Hall et al., 2011, Hall & Magner, 2011). The first volume addresses concepts of use, initial system requirements, architecture, and AeroMACS design considerations. The second volume describes AeroMACS prototype performance evaluation and provides final recommendations.

6.1 AeroMACS prototype design considerations

The AeroMACS prototype within the NASA-CLE CNS Test Bed is designed to implement the proposed AeroMACS features that are required to provide modern secure broadband wireless data communications at operational airports across the NAS. An essential element in the design and deployment of an AeroMACS network is a comprehensive radio frequency (RF) or physical layer (PHY) design.

An accurate RF design ensures that the deployed wireless network provides the necessary coverage, capacity, and reliability, with minimal interference, that satisfies the service requirements. Although it is possible to gauge the performance of radio links through theoretical means, real-life deployments must take into account variables from the environment to achieve optimal performance and minimize coverage holes and RF co-channel interference.

Figure 8 illustrates a top-level process for designing AeroMACS networks. The network design process begins with a physical site survey to gather information about the deployment location. A site survey provides an opportunity to validate any topography mapping information that may be available. It is also used to identify suitable installation locations for AeroMACS equipment. A site survey also provides input to the next three phases of the RF design process—coverage model, spectrum analysis, and capacity analysis.
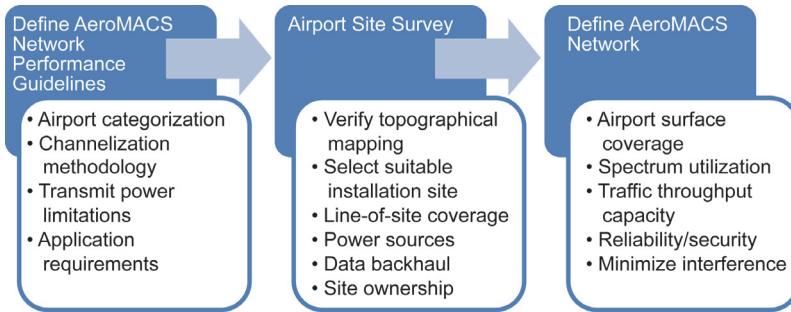
Fig. 8. AeroMACS network design process.

6.1.1 Coverage model

The coverage model requires a map of the site along with coordinates of potential locations for base stations (BSs) and user terminal subscriber stations (SSs). The coverage model must account for the impact of the environment on RF transmissions, including the effects of the topography, physical obstructions, and foliage. These effects introduce propagation loss and delays that have been cataloged in reference models. In addition, clutter models or obstruction densities are also modeled in this phase. Clutter models represent the density of obstructions in the deployment site. Typical options include rural, urban, and suburban clutter models. An airport surface with its relatively open runways and taxi areas and congested terminal areas will require a combination of the three models.

In addition to considerations of site topology and propagation delays, general parameters of the AeroMACS solution must be identified. Notable parameters include BS and SS transmit/receive power, antenna gains, feeder losses, BS and SS heights, and orthogonal-frequency-division multiple access (OFDMA) radio-access-related parameters. In addition, the following are the relevant system design parameters:

- Fade margin allocation for required link reliability
- Antenna gain and polarization diversity
- Co-channel interference margin
- Modulation and error correction
- Uplink to downlink transmit ratio (UL/DL ratio) for TDD mode
- Data throughput capacity requirements, including excess capacity margin
- BS and SS receiver noise figures
- Maximum BS and SS output/input power

Finally, a link budget must be calculated that specifies the maximum path loss between BS and SS locations. Receiver sensitivity for supported modulation schemes can be obtained from the BS and SS vendor data sheets. Characteristics of the BS and SS and information about the placement and types of antennas are used to generate an accurate coverage map.

6.1.2 Spectrum analysis

The spectrum analysis phase of the network design involves analysis of a potential site for interference. This includes both interference into the proposed AeroMACS and the potential for AeroMACS to interfere with co-allocated services. Interferers can include emissions at the fundamental frequency plus transmitter harmonics and inter-modulation emissions.

Proper analysis involves measurement of the maximum transmitter signal levels to determine how much energy is present across the surveyed RF band of interest. In the case of AeroMACS, C-band is the band of interest. The spectrum analysis can be conducted at ground level, but it is typically conducted from elevated locations including rooftops and tower sites at least 16 meters high.

6.1.3 Capacity analysis

Capacity analysis involves calculating how much traffic can be supported given the UL/DL ratio and the anticipated traffic patterns with the specified bandwidth and modulation scheme. The parameters used for capacity calculations include:

- TDD UL/DL ratio
- Modes of operation
- Channel bandwidth
- Subcarrier allocation scheme
- Transmit to receive guard ratio timing

The theoretical PHY throughput per modulation scheme can be calculated using the following formula (Upase et al., 2007):

$$R_b = R_s\, MC / R_r \tag{1}$$

Where:

M modulation gain (2 for quadrature phase-shift keying (QPSK), 4 for 16-quadrature amplitude modulation (QAM), and 6 for 64-QAM)

C coding rate (1/2, 3/4, 2/3, or 5/6)

R_r repetition rate (1, 2, 4, or 6)

R_b bit rate

R_s symbol rate

Equation (1) accounts for the AeroMACS modulation OFDMA pilot overhead but does not account for the signaling overhead. The signaling overhead depends on the number of active connections and the service types used. Studies have found that signaling overhead may vary from 4 to 10 percent of physical layer (PHY) throughput. Estimates of capacity using RF design tools take into consideration the impact of multiple-input, multiple-output (MIMO) antenna schemas to enhance coverage and/or capacity.

Although theoretical and software-based tools provide a baseline for determining the capacity of an AeroMACS network, it will be necessary to make minor adjustments once the network has been implemented. Such optimization involves selecting appropriate network parameters that will support the Quality of Service (QoS) requirements. A thorough mobile test drive throughout the deployed network is the final step for collecting network performance data for analysis and optimization.

6.2 AeroMACS prototype network architecture

The AeroMACS prototype was architected according to the reference network model developed by the WiMAX Forum® Network Working Group (WiMAX Forum®, 2011b). The reference network model is designed to enable interoperability of vendor equipment and to provide a structure for the deployment of new systems. The architecture is Internet Protocol (IP) based, meaning it relies on IP addressing to provide secure connectivity between users and access to common services. All WiMAX™ reference model elements are used to

implement the AeroMACS prototype network. These include: mobile SSs, stationary BSs, the access services network (ASN) function, and the connectivity services network (CSN) functions. The CSN functions include authentication, authorization, and accounting (AAA) and network management system (NMS).

6.3 AeroMACS prototype implementation

The AeroMACS prototype uses commercial WiMAX from the Alvarion® BreezeMAX® product line. Two BSs are included in the AeroMACS prototype to provide coverage redundancy and at least two opportunities for a mobile SS unit to link with a BS. One BS is located on NASA Glenn property and the second BS is on the CLE airport. Multiple base transceiver station (BTS) sectors are implemented at each BS to increase coverage, link sensitivity, and data capacity. The network includes ASN–gateway CSN functions to provide quality of service (QoS) control, user authentication and authorization for security, and mobility handoff between BSs and adjacent BTS sectors.

Many of the decisions about network layout for implementing the AeroMACS prototype in the NASA-CLE CNS Test Bed were driven by the need to use existing mounting structures for the BS and fixed SS sites, the desire to integrate with pre-existing test bed MLAT sensor sites, and the fact that the AeroMACS prototype is intended for experimental and demonstration purposes. As such, it does not interact with live airport operations and is not optimally configured for use as an operational system.

Figure 9 shows the placement of the two AeroMACS prototype BS sites in the NASA–CLE CNS Test Bed. BS-1, mounted on the tower adjacent to NASA Glenn's Flight Research Building (B4) hangar office, has two BTS sectors that are directed at 55° and 200° azimuth from true north. These are mounted 20m above ground level as shown in the upper-left inset photograph in Figure 9. BS-2, located on the roof of the Aircraft Rescue and Firefighting (ARFF) building located on CLE airport property, has three BTS coverage sectors directed 45°, 185°, and 295° from true north. The antenna mast and AeroMACS outdoor units (ODUs) are shown in the lower-right inset photograph in Figure 9. The ODUs are mounted to the mast on standoff arms to increase separation and RF isolation between units to thereby decrease the potential for in-band interference.
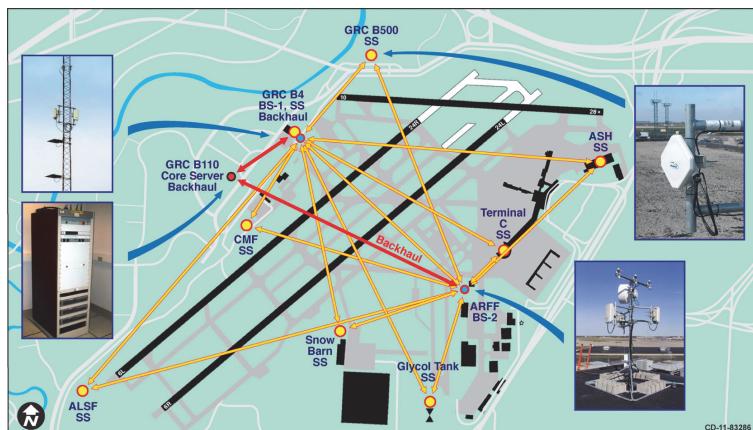

Fig. 9. NASA–CLE CNS Test Bed showing locations of the AeroMACS prototype base stations, fixed subscriber stations, microwave backhauls, and core server.

GPS outdoor units are mounted above each BTS ODU. Two are mounted on the tower at BS-1 at 3m above each BTS ODU, and three are mounted at BS-2, one for each BTS ODU. The GPS ODUs support precise timing and transmit/receive synchronization between BTS sectors. An option to reduce cost is to locate one GPS ODU per BS site and "chain" the GPS timing signal between ODUs. The coverage area of each BTS sector is 90° in azimuth as determined by the –3-dB pattern roll-off of the BTS sector antenna. These sector-coverage placements provide a high-degree of redundant coverage across the desired coverage area, including the runways, most of the taxiways, and much of the ramp areas.

Data from each BS site is transported to the core server using wireless backhaul links that operate in a licensed 11-GHz commercial band. A pair of these microwave radios is used on the roof of NASA Glenn's Space Experiments Building (B110) in full duplex operation between each BS site and the core CSN servers located in B110.

Figure 9 also shows the placement of SSs at eight fixed sites. Each of these sites was chosen for its co-location with MLAT surveillance sensors that were previously installed by the Sensis Corporation in NASA-CLE CNS Test Bed through a cooperative agreement with NASA Glenn. The Sensis MLAT sensors in the test bed were previously interconnected in a fixed wireless mesh network configuration that was based on the IEEE 802.11 standard (DeHart & Budinger, 2008). In the AeroMACS prototype, data from each MLAT sensor is transmitted wirelessly over the IEEE 802.16-based network to a central surveillance data processor. These MLAT sites are representative examples of fixed CNS infrastructure that AeroMACS can support within a mobile communications network on the airport surface.

A weatherproof enclosure is mounted near each SS as shown in Figure 10 only to support testing of the AeroMACS network. An operational AeroMACS network does not require such support equipment. The photograph shows the electronics equipment partially wired during construction.

Fig. 10. Electronics equipment supporting prototype subscriber stations during testing.

Each enclosure includes a single-board computer, a managed Ethernet switch, and power supplies to enable performance testing and applications demonstrations. The single-board computer hosts a Linux operating system and Ixia IxChariot® software for network performance tests. The IxChariot® software generates test data streams that are used to test

communication link capabilities. A test console is located at the core server in NASA Glenn B110 to coordinate the execution of tests, collect IxChariot® test results through the network, and compute statistics of network performance. Existing airport sensors, such as the MLAT surveillance remote units, can be connected as live data sources in place of, or in addition to, the IxChariot® software test data streams. A port on the managed switch is the interface for IP-based sensors such as the Sensis MLAT sensors.

6.3.1 Emulation of surface vehicle mobility

The range of vehicles that may use an operational AeroMACS network for communications vary from slow service vehicles (that mostly operate in terminal areas) to aircraft (that enter the network at relatively high-speed shortly after landing). The mobile environment for an arriving aircraft will transition from the mostly open, low-multipath conditions of the movement area to the terminal and gate area where multipath will increase but ground speeds are lower. The propagation environment will transition back to high speeds in mostly open areas as the aircraft departs the terminal gate and taxis for takeoff.

The NASA Aeronautical Research Vehicle (ARV), shown in Figure 11, was modified for use as a mobile AeroMACS SS under the various conditions expected for the airport surface environment.

Fig. 11. AeroMACS mobile SS logical network superimposed on NASA Glenn Aeronautical Research Vehicle and roof-mounted omni-directional antennas for mobility testing.

An AeroMACS SS unit and two antennas were mounted on the roof of the ARV to support mobile AeroMACS tests. An aluminum plate was used to form a ground plane for the two AeroMACS antennas as shown in Figure 11. A mobile AeroMACS SS unit, modified with RF connectors for attachment of external antennas, was mounted beneath the aluminum plate. The onmidirectional antennas used in the mobility tests are model SWA2459/360/20/V_2 from HUBER+SUHNER. These antennas exhibit constant gain of +8 dBi in ground plane directions. The gain pattern peaks toward the horizon because of the antenna orientation on the ARV.

Several fixed performance experiments and a set of initial mobility performance tests have been conducted successfully within the NASA-CLE AeroMACS prototype. Initial tests have explored the unique propagation conditions of an airport surface environment at C-band frequencies and the effects of AeroMACS profile parameter settings. Data throughput and packet integrity are measured for 5-MHz channel bandwidths for both stationary and mobile SSs.

The mobile SS integrated into the ARV was used to measure the performance at representative speeds of vehicles on the surface. The ARV was also used to verify the performance requirements to provide AeroMACS services on runways, taxiways, ramp areas, and gates. Initial mobility testing explored the transmit power required to maintain a minimum level of link performance for mobile SSs at vehicle speeds up to 50 knots using both single antenna and multiple-input multiple-output (MIMO) antenna diversity. Findings and recommendations are described in the following sections.

6.3.2 Runway drive tests

The first in a two-month series of mobile AeroMACS drive tests in the U.S. was conducted using the NASA ARV at the CLE airport on runway 24L/6R on 12 October 2010. Runway 24L is approximately 3 km in length, providing an opportunity to test AeroMACS air link ranges up to approximately 1.71 km. In addition to reduced signal strength caused by increased range, signal strengths also vary because of the antenna gain rolloff of the sectorized BS antenna. The positions of BS-1 and BS-2 relative to runway 24L/6R are marked in Figure 12.

The sector antenna pointing directions are indicated by white arrows for the BTS sectors (two for BS-1 and three for BS-2). The BTS sector antennas have a 90° half-power (–3 dB) beam width. The approximate–3 dB boundaries are indicated in Figure 12 with dashed lines for the two sectors used most often in these tests. The ARV travelling along runway 24L in the southwest (SW) direction experienced varying signal levels from a combined effect of changing range and BS sector antenna gain variation as the aspect angle changes.

Drive speed was nominally 40 kt. Tests were conducted with the mobile SS antenna system in MIMO and SISO modes. Network performance was evaluated by generation of bi-directional traffic using network test software. AeroMACS radio and network parameters were set up according to Table 3 for these tests.
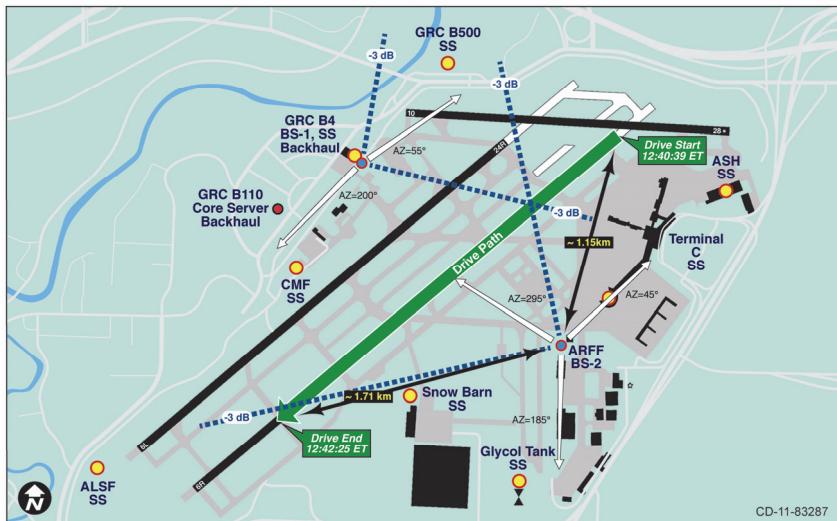

Fig. 12. AeroMACS mobility drive test on Runway 24L.

A plot of DL (BS to SS traffic direction) throughput during an ARV drive test along runway 24L in the SW direction is shown in Figure 13. The antenna configuration for this test uses 2 transmit antennas for the BS and 2 receive antennas for the mobile SS that is mounted on the ARV. This DL antenna configuration is referred to as 2x2 MIMO Matrix A (Space Time Block Coding). The highest average throughput expected on DL in a 5 MHz channel is 7.5 Mbps, which was achieved mid-way through the drive test. This corresponds to QAM64 modulation, the highest-order modulation supported by the standard.

AeroMACS Parameter	Setting				
AAA server	Enabled				
PKMv2, EAP-TTLS security	Enabled				
AES-128 air link encryption	Enabled				
Maximum transmission unit size	1440 bytes				
DL/UL ratio	60/40				
HARQ	Enabled				
MIMO	Matrix A mode enabled				
Channel bandwidth	5 MHz				
Quality of service (QoS)	Best effort				
BTS number	BTS1-1	BTS1-2	BTS2-1	BTS2-2	BTS2-3
BTS center frequencies, MHz	5100	5140	5130	5120	5110
BTS Tx power, dBm	21	21	21	21	21

Table 3. AeroMACS parameter settings for runway drive tests.

The IEEE 802.16 standard specifies an adaptive modulation feature for the SS that adapts the modulation rate according to link conditions with the goal of adapting data throughput rate to the highest level supportable by current link conditions. Test traffic throughput was reduced at the start and finish of the drive path, consistent with reduced modulation rate because of added propagation loss and BS sector antenna gain roll-off.
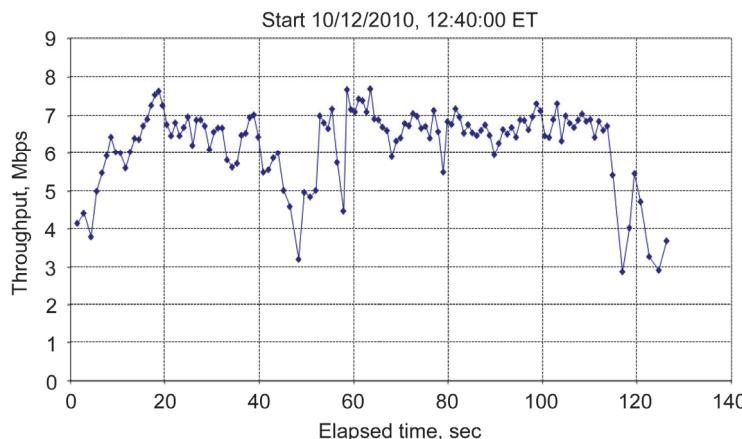

Fig. 13. Downlink throughput in MIMO antenna mode during drive test on Runway 24L.

The plot in Figure 14 compares the throughput performance of MIMO and SISO antenna configurations along the same drive path and for service provided by the same sector of BS-2 in both cases. A comparison of MIMO versus SISO throughput along the drive path shows that the MIMO antenna configuration achieved greater minimum and average throughput rates. Throughput averaged over the drive tests for MIMO and SISO antenna configurations are compared numerically in Table 4.
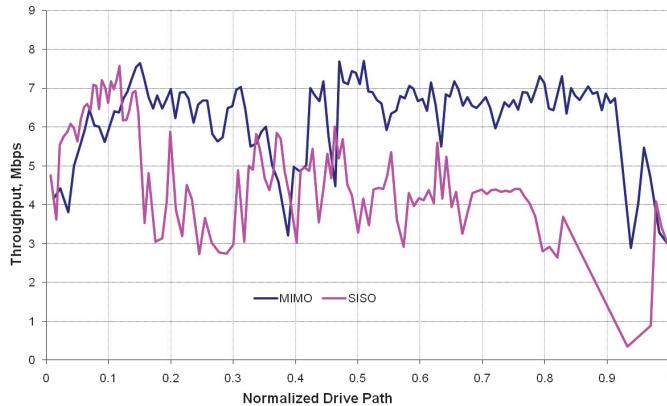

Fig. 14. Comparison of downlink throughput in MIMO and SISO antenna modes.

Test Time	Antenna Mode	Throughput Average, Mbps	Throughput Minimum, Mbps	Throughput Maximum, Mbps
1640 GMT	MIMO	5.13	2.70	7.70
1708 GMT	SISO	3.89	0.35	7.57

Table 4. MIMO and SISO mobile antenna configuration throughput comparison.

The runway 24L tests provide an initial assessment of mobile station antenna configuration impact on performance. The MIMO drive tests provide information on a unique antenna combination. The BTS antenna configuration is 2x2 MIMO in the AeroMACS prototype. Two antennas are arranged orthogonally to provide dual 45° slant polarization relative to the ground horizon. This test configuration represents a realistic scenario where BTS antennas use 45° slant-polarization to be compact and the SS antennas are spatially separated on a ground plane as they will be for an aircraft installation.

6.3.3 Base station transmit power requirements

BS transmit power level requirements were evaluated through a series of drive tests with the mobile ARV SS. Transmit power levels must be chosen to provide communication coverage across an airport surface while also minimizing potential interference to co-allocated users of the AM(R)S 5091- to 5150-MHz band. The survey of BS signal strength across the airport surface was used to assess whether adequate signal is radiated by the BSs. The signal strength survey was completed with a BS transmit power of +21 dBm (125 mW) to provide a benchmark level.

Drive tests along runway 24L were further analyzed for their implications for BTS transmit power requirements to provide an initial assessment of transmit power requirements. Additional analysis should be completed with future test data under additional drive test conditions. The ARV drive path driven at 1640 GMT is shown in Figure 12 with link distances shown from BS-2 to the start and end positions for the drive. The end of the drive provides the longest path distance of 1.71 km.

RSSI is a function of the link distance and BTS sector antenna gain. Real-time RSSI values from the ARV SS can be read periodically. These RSSI values are plotted in Figure 15 and are overlaid with data throughput measurements. Correlation between SS RSSI and throughput rate can be observed with higher RSSI readings (less negative) generally yielding higher throughput rate. The Yellowfin™ receiver provides another method of RSSI measurement. The Yellowfin™ instrument is programmed to scan through the AeroMACS frequency range searching for valid BS transmissions. RSSI is recorded with reference to the BS center frequency when a valid BS transmission is detected. BS transmissions are received through a 0-dBi gain antenna mounted on the roof of the ARV.

The ARV SS maintained service from the same BTS sector throughout the drive test shown in Figure 12. RSSI values recorded by the Yellowfin™ at the BTS2-3 center frequency of 5100-MHz are also plotted in Figure 15. Again, a correlation can be observed between Yellowfin™ and ARV SS measured RSSI and throughput rate derived by IxChariot®. Lower RSSI readings from the Yellowfin™ compared to the SS readings can be attributed to its lower receive antenna gain of 0 dBi compared to 8 dBi for the ARV antenna.

A few interesting performance characteristics can be observed in Figure 15 as follows:
1. Throughput rate was reduced as expected at the drive path start and end where lower signal strength occurred because of increased link path loss and decreased BTS sector antenna gain.
2. DL throughput reached a rate of 7.5 Mbps, the highest rate expected for a 5 MHz channel bandwidth, 60/40% TDD ratio, and MIMO Matrix A antenna configuration.
3. RSSI readings from the ARV SS and the Yellowfin™ decreased and hence the throughput rate decreased unexpectedly from 20% to 50% of the drive path. The cause of this reduced RSSI is unknown; it might be caused by an unwanted variation the BTS sector antenna pattern.
4. A minimum throughput rate of 3 Mbps was maintained over the length of Runway 24L. This included a maximum link path of 1.71 km at the –3 dB BTS sector pattern.
5. Link connectivity was maintained at vehicle speeds of at least 40 kt.

The operating conditions of the NASA Glenn AeroMACS prototype in Cleveland provided a DL throughput rate of at least 3 Mbps for a range of approximately 1.71 km for the following conditions:
- Clear line of sight from BS2 to ARV SS on runway 24L
- BTS sector transmit power: +21 dBm (125 mW) per MIMO channel
- BTS sector: 2x2 MIMO, mode A
- ARV SS: 2x1 MIMO, mode A
- BTS sector antenna gain: +16 dBi
- ARV SS antenna gain: +8 dBi

This test has established that a reasonable traffic throughput and range can be established with 125-mW BTS transmitter power under benign link conditions. Additional tests and analysis need to be completed to assure that this power level supports links into areas of higher signal multipath and NLOS conditions.
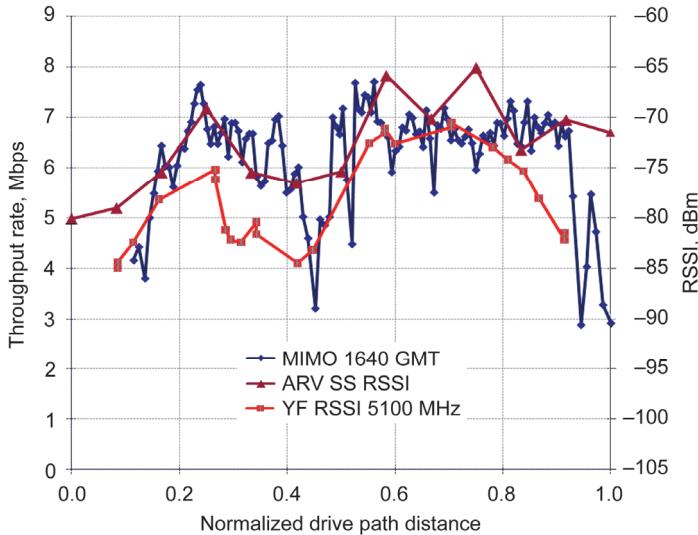
Fig. 15. Runway 24L drive test RSSI and throughput.

7. Conclusion

The ICAO approved concept for a broadband wireless mobile communications network to enhance safety and regularity of flight based on the IEEEs 802.16 standard is being realized through AeroMACS. An international standard is being pursued through collaboration between RTCA in the U.S. and EUROCAE in Europe. A wide variety of mobile and fixed applications are envisioned as candidates for AeroMACS in the U.S. The FAA, NASA Glenn and ITT have developed the world's first AeroMACS prototype in Cleveland, Ohio. Experimental measurements and mobility performance data from the prototype are being use to validate parameters of the WiMAX™ profile for AeroMACS. Further research and experimentation via the prototype and AeroMACS-equipped research aircraft will enable recommendations on total AeroMACS radiated power limits to avoid interference with collocated services, and potential performance and operational improvements from the use of MIMO antenna configurations. AeroMACS is the first component of the FCI expected to realize the ANC-11 vision for global harmonization of A/G communications.

8. Acknowledgment

The authors wish to acknowledge several people who contributed to the success of the FCS and the development of AeroMACS over the past 8 years. Mr. Brent Phillips served as the FAA co-sponsor for the FCS, NASA's AeroMACS research and development, and U.S. advocacy within the FAA and ICAO. Dr. Nikolas Fistas, EUROCONTROL, collaborated with the U.S. team from the beginning of AP-17 through joint development of standards via RTCA and EUROCAE. Mr. Rafael Apaza, FAA, provided valuable recommendations and practical considerations for design, development, and application of AeroMACS. Mr. Michael Biggs, FAA, and Dr. Izabela Gheorghisor, MITRE CAASD, provided guidance on the proper use of frequency spectrum, potential AeroMACS applications, and modeling for

9. References

Apaza, R. (2004). Wireless Communications for Airport Surface: An Evaluation of Candidate Wireless Technologies, *Proceedings of the 10th Ka and Broadband Communications Conference*, Vicenza, Italy, 30 September - 2 October 2004.

Apaza, R. (2010). User Services & Applications Survey (USAS) Ad-Hoc Working Group Report, *Joint Meeting of RTCA SC-223 and EUROCAE WG-82*, Brussels, Belgium 28-30 September 2010.

Biggs (2008). Outcome of the 2007 World Radiocommunication Conference (WRC-07), *Proceedings of ICNS Conference 2008,* Bethesda, Maryland, 5-7 May 2008.

Budinger, J., Hall, W., Wilson, J., Dimond, R., Apaza, R., & Phillips, B., (2010), FAA/NASA Aeronautical Mobile Airport Communications System (AeroMACS) Development Status, *16th Meeting of ICAO Aeronautical Communications Panel Working Group M,* ACP-WGM16/WP-17, Paris, France, 17-19 May 2010.

DeHart & Budinger (2008). Next Gen Airport Surface Wireless Network: Research Plan & Test Bed Performance, *Proceedings of I-CNS Conference 2008*, Bethesda, Maryland, 5-7 May 2008.

FAA (2011). Next Generation Air Transportation System (NextGen), (March 2011), Available from < http://www.faa.gov/nextgen/>

Fistas, N., Phillips, B., & Budinger, J., (2007) Action Plan 17 Future Communications Study - Final Conclusions and Recommendations Report, *(*November 2007), Retrieved from http://acast.grc.nasa.gov/media/Future_Communications_Study-Action_Plan_ 17_DASC_2007_Fistas_Phillips_Budinger.pdf>

Gheorghisor, I., (2008), Spectral Requirements of ANLE Networks for the Airport Surface, MITRE CAASD, MITRE Product MP080109R1, July 2008.

Gheorghisor, I., Hoh, Y., & Leu, A., (2009), Analysis of ANLE Compatibility with MSS Feeder Links, MITRE Technical Report MTR090458, December 2009.

Gheorghisor, I., Hoh, Y., & Leu, A., (2011), Compatibility of Airport Wireless Broadband Networks With Satellite Links in the 5091-5150 MHz Band, MITRE

CAASD, *Proceedings from ICNS 2011 Conference,* Herndon Virginia, 10-12 May 2011.

Gilbert, T., Dyer, G., Henriksen, S., Berger, J., Jin, J., & Boci, T., (2006) Identification of Technologies for Provision of Future Aeronautical Communications*,* ITT Industries, NASA/CR—2006-214451, October 2006.

Gilbert, T., Jin, J., Berger, J., & Henriksen, S., (2008), Future Aeronautical Communication Infrastructure Technology Investigation*,* ITT Industries, NASA/CR—2008-215144, April 2008.

Hall, E., Isaacs, J., Henriksen, S. & Zelkin, N., (2011), C-Band Airport Surface Communications System Standards Development, Phase II Final Report, Volume 1 Concepts of Use, Initial System Requirements, Architecture, and AeroMACS Design Considerations, NASA/CR—2011-216997-VOL1, April 2011.

Hall, E. & Magner J. (2011), C-Band Airport Surface Communications System Standards Development, Phase II Final Report, Volume 2 Test Bed Performance Evaluation and Final AeroMACS Recommendations, NASA/CR—2011-216997-VOL2, April 2011.

ICAO ANC (2003). Report of Committee B to the Conference on Agenda Item 7, *Proceedings of Eleventh Air Navigation Conference*, AN-Conf/11-WP/202, Montréal, Quebec, Canada, 22 September - 3 October 2003.

ICAO WGC (2006). Future Communications Study Working Papers and Information Papers*, ICAO Aeronautical Communications Panel (ACP) Working Group Communications* Meetings Reports (WGC-8, WGC-9, WGC-10, WGC-11), Available from: < http://www.icao.int/anb/panels/acp/wgmeetinglist.cfm?WGID=2 >

ICAO COCR (2007). Communications Operating Concept and Requirements for the Future Radio System (COCR) Version 2.0, *ICAO Aeronautical Communications Panel (ACP) Repository*, May 2007, Available from: <http://www.eurocontrol.int/ communications/gallery/content/public/documents/COCR%20V2.0.pdf/>

ICAO WGW (2008). Report of Agenda Item 1: Finalization of the Future Communications Study, *ICAO Aeronautical Communications Panel (ACP) Working Group of the Whole Second Meeting,* Montreal, Canada, 21-25 April 2008.

IEEE (2009). IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, IEEE Standards Association, Available from < http://standards.ieee.org/about/get/802/802.16.html>

ITU-R (2007). Initial estimate of new aviation AM(R)S spectrum requirements, Report ITU-R M.2120, Article E 31647, October 2007, Available from: < http:// www.itu.int/pub/R-REP-M.2120-2007/en>

Kerczewski, R., (2006), Bandwidth Requirements and Channelization Approaches for Future Airport Surface Communications in the 5091-5150 MHz Band*,* ACP-WGF15/WP23, Cairo Egypt, 7 – 13 June 2006.

Matolak, D., (2007), Wireless Channel Characterization in the 5 GHz Microwave Landing System Extension Band for Airport Surface Areas*,* Ohio University, NASA Grant NNC04GB45G, NASA/CR—2007-214456, May 2006.

Phillips, B., Pouzet, & J., Budinger, J., (2007), Future Communication Study Technology Recommendations - Action Plan 17 Final Conclusions and Recommendations Report, *ICAO Aeronautical Communications Panel Working Group Technology*, Working Paper ACPWGT/1-WP/06_AP17, Montreal, Canada, 2-5 October 2007.

RTCA SC-223 (2011). Special Committee SC-223, Airport Surface Wireless Communications, Available at <http://www.rtca.org/comm/Committee.cfm?id=133>

Sensis (2011). Airport Surface Detection Equipment, Model X (ASDE-X), Sensis Corporation, April 2011, Available from < http://www.sensis.com/docs/128/>

SESAR (2011). Single European Sky ATM Research, (March 2011), Available from http://ec.europa.eu/transport/air/sesar/sesar_en.htm/>

Upase et al. (2007). Radio Network Dimensioning and Planning for WiMAX Networks, Fujitsu Scientific & Technical Journal, vol. 43, no. 4, 2007, pp. 435–450, Tokyo Japan.

Wargo, C. & Apaza, R., (2011), Application Survey for the Future AeroMACS, *Proceedings from ICNS 2011 Conference*, Herndon Virginia, 10-12 May 2011.

Wilson, J. & Kerczewski, R., (2011), Interference Analysis for an Aeronautical Mobile Airport Communications System, *Proceedings of IEEE 2011 Aerospace Conference*, Big Sky, Montana, 5-12 March 2011.

WiMAX Forum® (2011a). Technical Specifications Air Interface T21, WiMAX Forum®, Available from <http://www.wimaxforum.org/resources/documents/ technical>

WiMAX Forum® (2011b). WiMAX Forum® Network Architecture - Architecture Tenets, Reference Model and Reference Points Part 0 - Release 1.0, WMF-T32-001-R010v05 - Available from http://www.wimaxforum.org/resources/documents/technical/ release >

Utilizing IEEE 802.16 for Aeronautical Communications

Max Ehammer, Thomas Gräupl and Elias Pschernig
University of Salzburg
Austria

1. Introduction

The future Air Traffic Management (ATM) concept shall be based on network centric operations, consequently on information sharing. In order to support such a vision not only a versatile and capable ground based communication network is necessary but also a network which includes the air to ground sub-networks which shall have sufficient capacity and capability. One such air to ground sub-network shall be established for the airport surface intended to be used by departing and arriving aircraft as well as by surface vehicles. This communication system is currently (2011) emerging and shall be called Aeronautical Mobile Airport Communications System (AeroMACS). AeroMACS shall be based on the IEEE 802.16-2009 standard (IEEE, 2009) and especially on the WiMAX Forum™ Mobile System Profile Specification rel1.0 v0.9 (WiMAX Forum, 2010). A draft profile has been developed and is being evaluated currently, e.g. in the EU research project SANDRA (SANDRA).

The IEEE 802.16-2009 standard (IEEE, 2009) specifies the air interface of combined fixed and mobile point to multipoint broadband wireless access systems with the possibility to support different services. The standard specifies the Medium Access Control (MAC) and the Physical (PHY) layer, where the MAC is capable to support multiple PHY specifications applicable to a specific operational environment. Figure 1 depicts the protocol reference model of the IEEE 802.16 standard. The Service Specific Convergence Sub-layer (CS) accepts higher layer data protocol units (PDUs) via the CS service access point (SAP). Thereby, the CS classifies each higher layer PDU according to available policies and maps each higher layer PDU to a so called service flow identifier. The IEEE 802.16 standard provides multiple CS specifications in order to provide interfaces for a variety of higher layer protocols. The MAC Common Part Sub-layer (CPS) provides the core functionality for data exchange via the wireless medium. A separate security sub-layer is also available. Generally, the IEEE 802.16-2009 standard provides a large amount of options. Thereby, different options may fit better for certain use cases than others. Due to the large amount of options it is merely impossible to be interoperable among different vendors based on the sole standard. Additional documentation and specification is necessary. This task has been conducted by the WiMAX Forum™. This group specifies so called "WiMAX profiles" where a selected set of options from the IEEE 802.16 standard is qualified for such a profile.

The WiMAX Forum™ has been established in June 2001 and is an industry led nonprofit organization. The purpose of the WiMAX forum™ is to certify compatibility and interoperability of broadband wireless products based on the IEEE 802.16 standard. In such a

way rapid introduction of technology and market competition shall be enforced. The WiMAX Forum™ has many members comprising the majority of operators and equipment vendors.
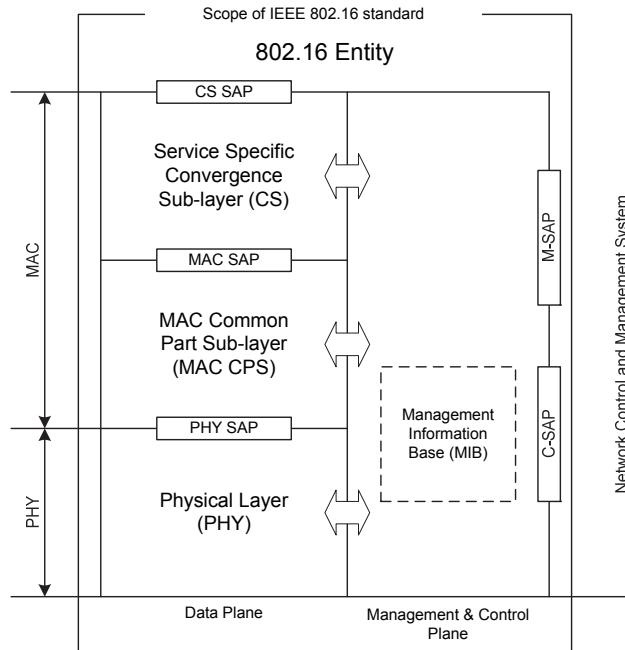

Fig. 1. The IEEE 802.16 protocol reference model.

The WiMAX Forum is organized into several working groups, one of them is the Technical Working Group (TWG), which develops technical product specifications and certification test suites for the air interface based on the OFDMA PHY. Such specifications are complementary to the IEEE 802.16 standards in order to achieve interoperability and certification of Mobile Stations and Base Stations conforming to the 802.16 standards. The TWG has produced a "Mobile System Profile Specification" which determines mandatory and optional functions.

Sub-chapter 2 gives an overview of selected AeroMACS profile items with some explanations. Sub-chapter 3 discusses the possibilities to run IPv6 over AeroMACS. Sub-chapter 4 summarizes the outcome of the data traffic load analysis conducted during the course of the SANDRA project. Sub-chapter 5 shows selected results of the medium access performance analysis. At the end concluding remarks finalize this chapter.

2. AeroMACS profile overview

The WiMAX Forum™ Mobile System Profile Specification (WiMAX, 2010) represents a subset of the IEEE 802.16 standard (IEEE, 2009). Currently (2011) a draft profile for AeroMACS is being specified through standardisation bodies such as EUROCAE and RTCA. This draft profile is further evaluated by members of the SESAR Joint Undertaking and by members of the SANDRA project (SANDRA, 2011). The tentative AeroMACS draft profile is further a subset of the WiMAX Forum™ Mobile System Profile Specification.

Within this sub-chapter an overview of the core functionalities related to data exchange are given.

2.1 Overview physical layer

The Physical Layer (PHY) of the AeroMACS system shall be based on the OFDMA Physical Layer specification of the IEEE 802.16 standard with a channel bandwidth of 5 MHz. Thereby, the frame length shall be 5 ms. As the PHY will be based on the "Common part TDD profile" (WiMAX 2009), the Downlink and Uplink portions can vary dependent on the system settings (cf. Figure 2).

The IEEE 802.16-2009 supports both Time Division Duplexing (TDD) and Frequency Division Duplexing (FDD) modes. However, AeroMACS shall be based on the TDD mode of operation. Reasons therefore are the dynamic allocation of Downlink (i.e. from Base Station to Mobile Station) and Uplink (i.e. from Mobile Station to Base Station) resources in order to efficiently support asymmetric Downlink (DL)/Uplink (UL) traffic, only a single channel is required which alleviates spectrum issues, and the TDD option is less complex.
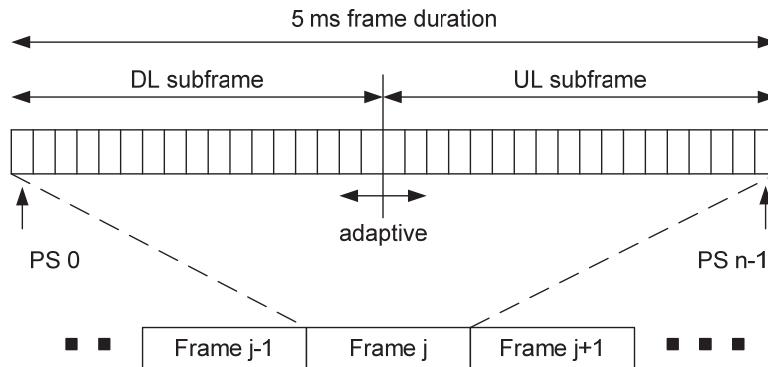

Fig. 2. AeroMACS frame with an adaptive DL/UL subframe width.

The DL subframe and the UL subframe consist of a number of OFDM symbols where a reasonable setting could be 29 OFDM symbols for the DL and 18 OFDM symbols for the UL. However, the individual setting is dependent on the service provider. Valid values can be taken from the WiMAX Forum™ Mobile System Profile Specification TDD Specific Part (WiMAX, 2009).

The standard supports multiple schemes for dividing the time and frequency resources among users, this may also be called sub-channelization. AeroMACS shall be based on the pseudo-random permutation for frequency diversity (i.e. PUSC). The available spectrum has to be utilized by the resource scheduler through an integer number of DL and UL slots, respectively. A slot is a logical $n \times m$ rectangle where n is the number of sub-carriers and m is the number of contiguous OFDM symbols. All slots, no matter which sub-channelization scheme is being used, contain 48 data symbols. Thereby, a DL slot consists of 2 OFDM symbols and 28 subcarriers. As the total usable amount of subcarriers is 420 for the DL, this results in 210 usable DL slots per 5 ms frame in the downlink direction (considering 28 OFDM symbols plus 1 OFDM symbol used for the DL Prefix). In contrast a UL slot consists of 3 OFDM symbols and 24 subcarriers. For the uplink direction the total usable amount of

subcarriers is 408, consequently there are 102 usable UL slots per 5 ms frame in the uplink direction (assuming 18 OFDM symbols).

Dependent on the coding and modulation scheme different throughput can be achieved. The modulation schemes are QPSK and 16 QAM for both directions as well as 64 QAM for the DL direction. 64 QAM is also being discussed as an option for the UL direction. Dependent on the robustness of the coding scheme different theoretical throughput values can be achieved.

	DL			UL		
	QPSK	16 QAM	64 QAM	QPSK	16 QAM	(64 QAM)
CC 1/2	48 bits	96 bits	144 bits	48 bits	96 bits	(144 bits)
CC 2/3	64 bits	128 bits	192 bits	64 bits	128 bits	(192 bits)
CC 3/4	72 bits	144 bits	216 bits	-	-	-
CC 5/6	80 bits	160 bits	240 bits	80 bits	160 bits	(240 bits)

Table 1. Data size per DL/UL slot with various modulation and coding schemes.

	DL (28 OFDM symbols)			UL (18 OFDM symbols)		
	QPSK	16 QAM	64 QAM	QPSK	16 QAM	(64 QAM)
CC 1/2	2,016 Mbit	3,859 Mbit	6,048 Mbit	0,979 Mbit	1,958 Mbit	(2,9 Mbit)
CC 2/3	2,572 Mbit	5,376 Mbit	8,064 Mbit	1,305 Mbit	2,611 Mbit	(3,9 Mbit)
CC 3/4	3,024 Mbit	6,048 Mbit	9,072 Mbit	-	-	-
CC 5/6	3,360 Mbit	6,720 Mbit	10,08 Mbit	1,632 Mbit	3,264 Mbit	(4,9 Mbit)

Table 2. Raw data rate in megabits per second considering a setting of 28 usable OFDM DL symbols and 18 usable OFDM UL symbols.

A broad range of combinations exists, however, most likely is a combination with robust coding (i.e. convolution code (CC) with rate 1/2) with modulation of QPSK or 16 QAM for the UL and 16 QAM or 64 QAM for the DL.

Each 5 ms AeroMACS frame starts with a DL Prefix which occupies one entire OFDM symbol. The Frame Control Header (FCH) follows immediately the DL Prefix and contains information about the following DL Map. The DL Map and the UL Map are important management elements which tell the Mobile Stations (MSs) how the upcoming frame is to be used to exchange either data or management information. The mentioned elements of DL Prefix, FCH, DL Map, and UL Map appear in each DL subframe. The UL direction needs to schedule ranging opportunities for Mobile Stations in order to keep synchronized with the Base Station and in order to request bandwidth if a MS needs to do so. The remaining bandwidth may be used to transmit user data.

2.2 Overview medium access control

The IEEE 802.16 standard specifies a point to point and connection oriented link, i.e. each Service Data Unit (SDU) received from an interfacing higher layer is mapped to a unique and unidirectional service flow with specific quality of service (QoS) parameters. Thereby, the interfacing higher layer can be one of several different types.

The MAC common part sub-layer operates in a point to multipoint environment. The Base Station (BS) is the only user of the Downlink (DL) resources, whereas the Mobile Stations have to share the Uplink (UL) resources. All MSs are able to receive DL transmissions. Based

on the Connection Identifier (CID) carried within the generic MAC header of each MAC PDU a MS is able to determine whether a MAC PDU is destined to it or not.

A central concept of the IEEE 802.16 standard is the usage of transport connections which allows the utilization of QoS at MAC level. Each service flow has specific QoS parameters initialized at connection setup. Thereby, different data delivery strategies can be utilized (e.g. best effort, polling, etc.).

At system initialization two pairs of management connections, namely the basic connection and the primary management connection, have to be established between the MS and the BS. A third management connection, the secondary management connection, may be established, too. However, such a connection is only mandatory for managed "subscriber stations". In certain circumstances especially if remote airport equipment is being used such a secondary management connection would probably make sense. However, the basic management connection shall be used to transmit short and time urgent MAC management messages while the primary management connection shall be used to exchange longer and delay tolerant MAC management messages.

2.3 Convergence sub-layer options

The convergence sub-layer (CS) in the case of the IEEE 802.16 standard specifies the interface towards higher layer protocols. The standard provides a variety of convergence sub-layer options in order to provide the possibility to interface with a versatile set of higher layer protocols. The principle functions of the convergence sub-layer are accepting and interpreting the higher layer protocol header and consequently the mapping of this information to a specific service flow. Additionally, header compression techniques or any other appropriate processing may be conducted by the convergence sub-layer protocol.

The AeroMACS draft profile foresees only the IP CS (support of IPv6), however, optionally also the Ethernet CS is supported. In principle the issue of the convergence sub-layer seems straight forward - either AeroMACS supports higher layer protocol A or higher layer protocol B. However, recalling the principle design issues of layered communication protocol architectures there may be issues as discussed below.

Figure 3 shows a generic view of a network reference model containing layers, interfaces, and protocols. Thereby, a layer of a network can be seen as an abstraction of a service or services to be provided to its higher layer. In such a way the implementation details are hidden from the service user, that is the higher layer. The higher layer accesses these services through a well defined interface (also known as service access point). The specification of clean interfaces provides the advantage to exchange completely different implementations of layers, assuming that the new implementation provides the same set of services as the old implementation did. Virtually, two communicating hosts are exchanging information on layer n using a layer n protocol. However, real communication takes place only via the physical transmission medium. Generally, a protocol specifies the rules and conventions to be used in a communication.

A list of protocols used by a certain system, one protocol per layer, is called a protocol stack. Services and protocols are distinct concepts, although they are frequently confused. A service is a set of primitives (operations) that a layer provides to the layer above it. The service defines what operations the layer is prepared to perform on behalf of its users, but it says nothing at all about how these operations are implemented. A service relates to an interface between two layers, with the lower layer being the service provider and the upper layer being the service user.
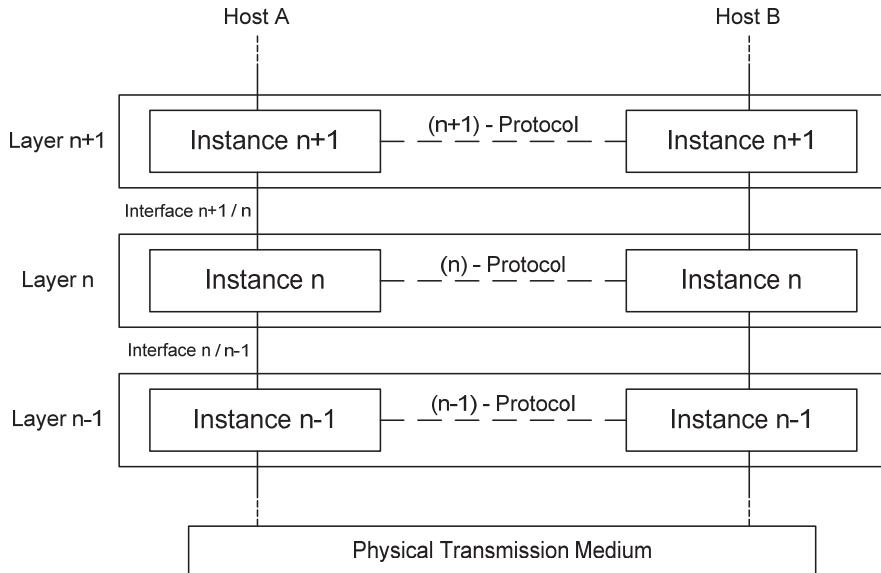
Fig. 3. A generic protocol stack.

A protocol, in contrast, is a set of rules governing the format and meaning of the frames, packets, or messages that are exchanged by the peer entities within a layer. Entities use protocols in order to implement their service definition. They are free to change their protocols at will, provided they do not change the service visible to their users. In this way, the service and the protocol are completely decoupled.

Considering the two options of the packet based convergence sub-layer, namely IP CS and Ethernet CS. The offered service differs from the IP point of view and may cause problems when considering IP over AeroMACS (c.f. Chapter 3).

2.4 MAC PDU formats

The IEEE 802.16 standard offers various options for fragmenting and reassembling MAC Service Data Units. Thereby, a MAC SDU may be of variable or fixed length. In the case of AeroMACS a variable length of MAC SDUs shall be allowed. Generally, a MAC Protocol Data Unit shall be of the form as depicted in Figure 4. Each MAC PDU starts with a fixed length header of 6 bytes (the generic MAC header). A MAC PDU typically contains payload and shall then be appended by a 4 bytes Cyclic Redundancy Checksum (CRC). The payload itself may further contain several sub-headers (SH). The fragmentation sub-header is used if an entire MAC SDU does not fit into a single MAC PDU. The packing sub-header is used if several MAC SDU are packed together into a single MAC PDU. Multiple MAC PDU may also be concatenated during a single burst.

If the MAC PDU does not contain payload data MAC header needs no CRC as the MAC header itself contains a header checksum. The generic MAC header contains the connection identifier (CID) of the connection with which the PDU is associated. The MAC PDU does not contain any source or destination address in its header. The tentative AeroMACS draft profile uses the same MAC PDU format specification as the WiMAX Forum (WMF) Mobile System specification (WiMAX Forum, 2010).

2.5 Automatic Repeat Request (ARQ)

Generally, Automatic Repeat Request (ARQ) protocols are used to synchronize data flows between sending and receiving entities. Thereby, the flow control procedure takes care that the data source is not overloading the data sink. Also erroneous data packets are indicated to the source (through negative acknowledgments).
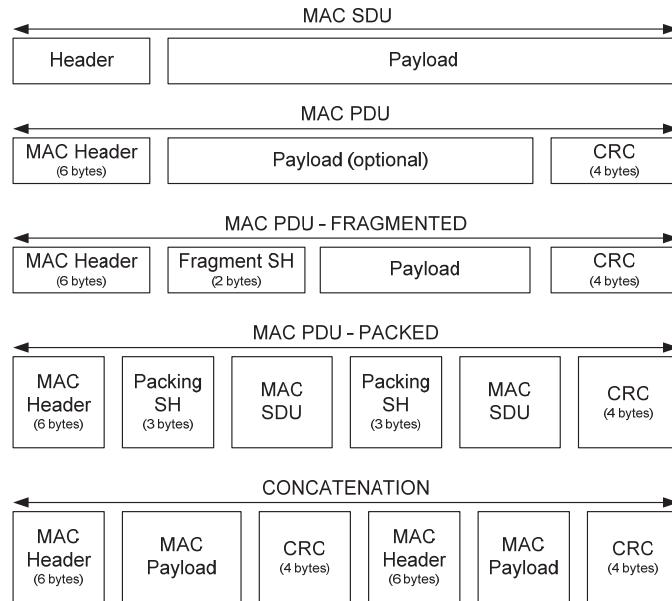

Fig. 4. Overview of different MAC PDU options.

The IEEE 802.16 standard offers four different types of ARQ, namely, go-back-n, selective-reject, and two combinations of go-back-n and selective-reject. Go-back-n may also be called as cumulative ARQ. An ARQ information element has at least a size of 4 bytes and at most of 12 bytes. The basic components of an ARQ information element are the connection identifier field and the block sequence number (BSN) field. The CID identifies the transport connection and the BSN is differently used dependent on the ARQ type.

The ARQ protocol of the IEEE 802.16 standard is based on ARQ blocks, which all have a size of ARQ_BLOCK_SIZE in bytes. An exception may only be for the last ARQ block of an SDU which may be smaller. Each incoming SDU from a higher layer is logically divided into a number of ARQ blocks. Thereby, each ARQ block gets a BSN. Compare Figure 5 which is showing an example with ARQ_BLOCK_SIZE set to 32 bytes and a sequence of three arriving SDU with a size of 90, 10, and 64 bytes.
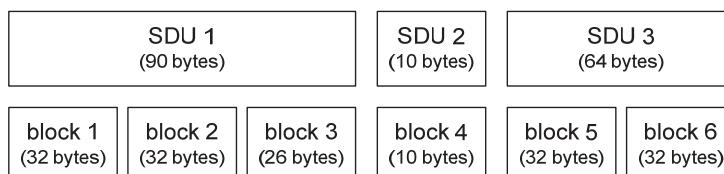

Fig. 5. Example of different SDU with an ARQ block size of 32 bytes.

The single blocks may be packed into one or several MAC PDUs. Using the above example all 6 ARQ blocks could be packed together to a single MAC PDU, however, the ARQ blocks could also be packed into 3 separate MAC PDU where each MAC PDU carries 2 ARQ blocks. ARQ blocks from the same SDU with consecutive block sequence numbers can be grouped together into an SDU fragment. Each fragment (or single ARQ block) is preceded by a packing sub-header (PSH) which carries the BSN of the first ARQ block and the length of the fragment in bytes. This allows the receiver to decode the ARQ blocks again. If the fragment length is not a multiple of ARQ_BLOCK_SIZE, it means the last block in the fragment is smaller. The complete MAC PDU for the example case above where all ARQ blocks are sent in a single MAC PDU could look like as depicted in Figure 6.
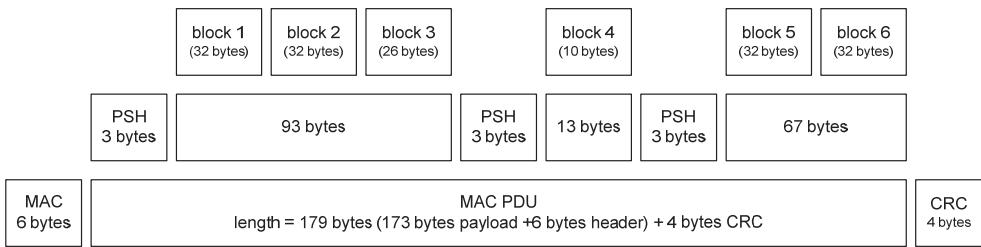

Fig. 6. Example MAC PDU - MAC packing.

Figure 6 continues the example from above where 3 MAC SDUs are packed into a single MAC PDU. Each MAC SDU (fragment) is prefixed by a PSH which has a length of 3 bytes. Additionally, the MAC PDU overhead accounts for 10 bytes - i.e. the generic MAC header with 6 bytes and the CRC with 4 bytes. This example of packing MAC SDU results in an overhead of 19 bytes for a payload of 164 bytes.

2.5.1 ARQ acknowledgment types

Acknowledgment type 0 (i.e. selective ACK entry) contains up to 4 fixed length acknowledgment maps. The length of such a map is 16 bits where each bit indicates whether a corresponding ARQ block has been received successfully (i.e. bit is set) or not (i.e. bit is not set). When using the selective ACK entry option the BSN corresponds to the first bit of the following acknowledgement map. It is important to realize that such an acknowledgement type is only applicable if more than or equal to 16 ARQ blocks have been received without prior sent acknowledgment.

Acknowledgment type 1 (i.e. cumulative ACK entry) uses the BSN to cumulatively acknowledge all ARQ blocks received. This acknowledgment type has a fixed size of 4 bytes. Acknowledgment type 2 (i.e. cumulative with selective ACK entry) is a combination of acknowledgment type 0 and type 1. In this case the BSN is interpreted as cumulative acknowledgment and the first bit of the following map is set - the remaining bits of the map can be used as in type 0.

Acknowledgement type 3 (i.e. cumulative with block sequence ACK entry) is a combination of type 1 and a series of sequence ACK maps. The BSN acknowledges all correctly received ARQ blocks cumulatively. The sequence ACK map contains either two sequences with a length given in 6 bits or three sequences with a length given in 4 bits. Thereby, each sequence specifies a number of consecutive BSN entries, with the first sequence starting at

the cumulative BSN plus one (which is always a negative acknowledgment; otherwise the cumulative BSN would be increased).

Figure 7 shows a hypothetical example of 32 contiguous ARQ blocks where some of them were received correctly and some of them were received erroneously. Erroneous blocks are marked with an X. The differences of the various acknowledgment types become obvious when applying each acknowledgment type separately to the same set of ARQ blocks. In this case type 0 (i.e. selective ACK entry) is capable to give feedback on every received ARQ block. The reason therefore is that the number of ARQ blocks fits exactly two acknowledgment maps of 16 bits. The type 1 acknowledgment (i.e. cumulative ACK entry) can only confirm the receipt of the first three correctly received ARQ blocks. The type 2 acknowledgment (i.e. cumulative and selective ACK entry) is capable of acknowledging only the first 18 ARQ blocks. The reason therefore is that selective ACK map sizes are fixed to a length of 16 bits, the remaining 14 ARQ blocks cannot be acknowledged by this method until erroneously received ARQ blocks are being retransmitted and received correctly. The type 3 acknowledgment (i.e. cumulative with block sequence ACK entry) uses sequences to acknowledge sequences of correctly or erroneously received ARQ blocks. If the option is used with 2 block sequences per sequence map (3a) only 20 ARQ blocks can be acknowledged. Using the option with 3 block sequences per sequence map (3b) acknowledges 27 ARQ blocks in this case. This example does not work well for the block sequence map option as the sequences are very short.
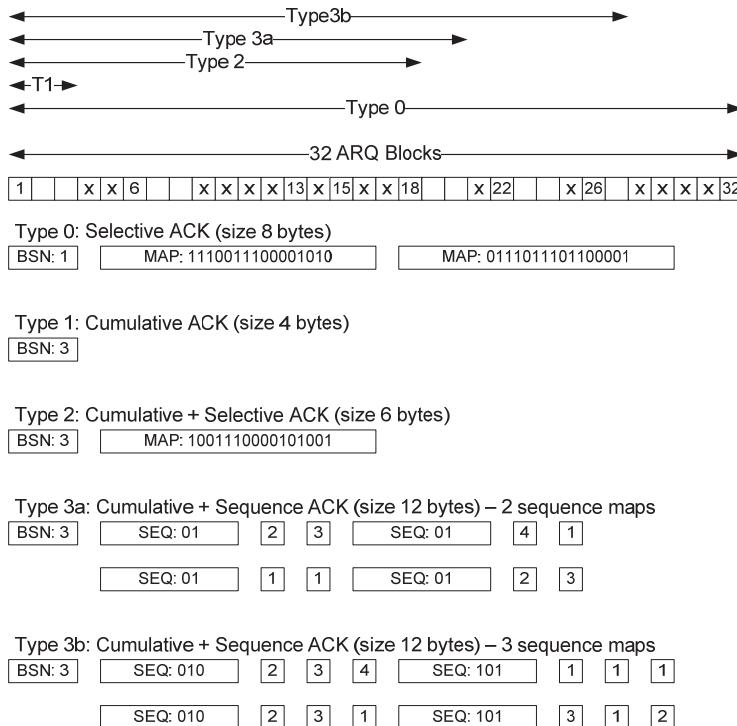

Fig. 7. Illustration of the capabilities of the different ARQ acknowledgment types.

This example illustrates the functionality of the different ARQ options which may be used by the AeroMACS profile. Each ARQ option has its advantages depending on the frequency acknowledgments are sent (e.g. after the receipt of each MAC PDU, or after a certain amount of time, etc.), the pattern of the occurred errors, or the computational complexity.

The WiMAX Forum™ Mobile System Specification requires ARQ acknowledgment types 1, type 2, and type 3 to be implemented. ARQ acknowledgment type 0 is optional. The AeroMACS profile intends to support the same set of acknowledgment options.

Each acknowledgment type has its advantages but is dependent on feedback intervals, error patterns, and computational complexity. The size of the ARQ block has an impact as well, if large ARQ blocks are used it is more unlikely to fill acknowledgment maps or acknowledgment sequence maps. The standard does not specify any strategy how and when ARQ acknowledgments shall be scheduled.

2.6 Qualtiy of Service (QoS)

Quality of Service in IEEE 802.16 is supported through the concept of unicast transport connections. These transport connections are called service flows, where each service flow utilizes a particular set of QoS parameters. The standard provides several QoS parameters to be adjusted; for instance maximum sustained traffic rate, maximum traffic burst, minimum reserved traffic rate, maximum latency, etc. - in principle latency, jitter, and throughput assurance.

Service flows are either provisioned or dynamically added by the Base Station or optionally by the Mobile Station. How to provision service flows is out of the scope of the IEEE 802.16 standard, consequently it is also not specified in the AeroMACS draft profile. Certain service flows may be added dynamically for instance after the network entry procedure. The standard provides options to create, change, and delete a service flow dynamically. Such a procedures can be either initiated by the Base Station or by the Mobile Station. The WiMAX mobile profile makes these options mandatory to be supported by the Base Station. The capability to dynamically create or change a service flow is optional for a Mobile Station, however, the deletion of a service flow is mandatory. The AeroMACS profile intends to support only the dynamic service flow creation, change, and deletion procedures to be initiated by the Base Station.

How these service flows are initiated and / or triggered is not specified by the AeroMACS profile. QoS parameters of ATC traffic flows shall probably be regulated while QoS parameters of AOC traffic flows may be provider dependent.

2.7 Scheduling & data delivery services

There are different possibilities to provide bandwidth to a Mobile Station, realized through a scheduling service. Uplink request and grant scheduling is performed by the Base Station in order to provide each Mobile Station with bandwidth for uplink transmissions or opportunities to request bandwidth. By specifying a scheduling type and its associated QoS parameters, the Base Station scheduler can anticipate the throughput and latency needs for the uplink traffic and provide polls and/or grants at the appropriate times. The different scheduling services are:

- Unsolicited Grant Service (UGS)
- real-time Polling Service (rtPS)
- non-real-time Polling Service (nrtPS)

- Best Effort (BE)

The unsolicited grant service (UGS) is intended for real-time applications which generate fixed-rate data. Among others QoS parameters such as tolerated jitter, minimum reserved traffic rate, maximum latency, and the unsolicited grant interval are defined. This means that a service flow with a data delivery service of UGS gets periodically UL resources assigned without requesting them each time.

The real-time Polling service (rtPS) is intended for real time applications with variable bit rates. Among others QoS parameters such as maximum latency, minimum reserved traffic rate, traffic priority, and the polling interval are defined. In this case the resource scheduler polls a Mobile Station regularly at fixed intervals. These polls may be used to request bandwidth.
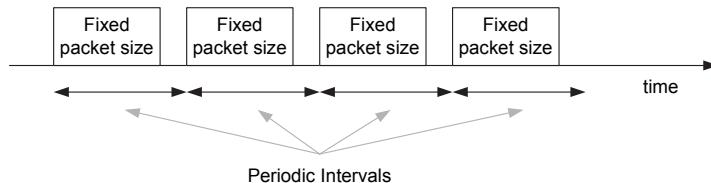

Fig. 8. The Unsolicited Grant Service (UGS) usable for uplink transmissions.
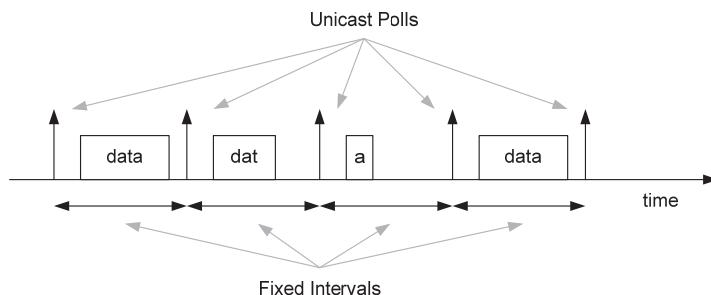
Fig. 9. The real-time Polling service (rtPS) usable for uplink transmissions.
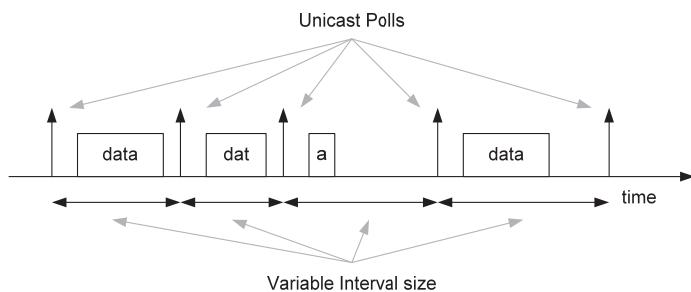
Fig. 10. The non-real-time Polling Service (nrtPS) usable for uplink transmissions.

The non-real-time Polling Service (nrtPS) is intended for applications which require guaranteed data rate but are insensitive to delays. QoS parameters such as minimum reserved traffic rate, maximum sustained traffic rate, and traffic priority are defined. In this case the unicast polls are issued at a variable interval length (dependent on the available resources). The polls may be used to request bandwidth.

The Best Effort (BE) service is intended for applications with no rate or delay requirements. In this case bandwidth request ranging opportunities are provided to transmit bandwidth request ranging codes. If a bandwidth request range code is successfully received by a Base Station it polls the associated Mobile Station.
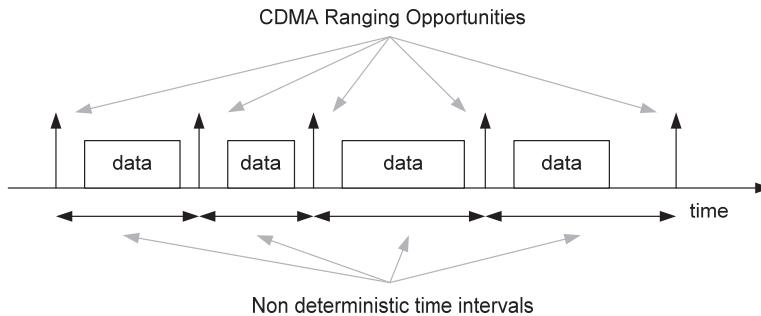

Fig. 11. The Best Effort (BE) service usable for uplink transmissions.

For the downlink direction similar QoS classes can be utilized. However, these are called slightly different but have comparable QoS parameters. The scheduler does not need to consider any polls or ranging opportunities for the downlink, though. The different QoS classes or data delivery services are:

- Unsolicited Grant Service (UGS)
- Real-Time Variable Rate (RT-VR) service
- Non-Real-Time Variable Rate (NRT-VR) service
- Best Effort (BE) service

What kind of QoS class a specific application or set of applications will require is dependent on the requirements. How the different data delivery services and scheduling strategies are implemented is not specified by the standard. Thus, they are vendor dependent. In any case the communication service provider has to assure that safety related communication is preferred over non safety related communication. It might be that a simple priority scheme with a best effort service is sufficient.

2.8 Request-grant mechanisms

A Mobile Station is required to support at least three different connections. That is, two management connections which are set up at network entry and one data bearer to transmit user data.

Every connection with a QoS service other than UGS needs to adapt its resource requirements. This is done through bandwidth requests. This is a mechanism where a Mobile Station indicates to the Base Station that it requires uplink resources. Bandwidth requests are either sent as standalone Bandwidth Request (BR) headers or as a Piggy Back Request (i.e. included in the Grant Management Sub-header (GMSH)).

Bandwidth Requests may be either incremental or aggregated. When a BS receives an incremental BR, it shall add the quantity of bandwidth requested to its current perception of the bandwidth needs of the connection. When the BS receives an aggregate BR, it shall replace its perception of the bandwidth needs of the connection with the quantity of bandwidth requests. Piggybacked bandwidth requests are always incremental.

The Base Station issues resource grants towards a Mobile Station based on the basic CID (i.e. basic management connection). This means that a Mobile Station is able to utilize the concept of bandwidth stealing where a certain amount of requested bandwidth for a specific QoS class may be utilized differently. However, the resource requests are based on the transport connection which requires bandwidth. If a Base Station polls a Mobile Station it typically assigns enough resources to issue a bandwidth request.

3. IPv6 over AeroMACS

The Network Working Group (NWG) of the WiMAX Forum™ has defined a network architecture for IEEE 802.16 sub-networks. Thereby, considering topics at layers above those defined by the 802 standards. The Internet Engineering Task Force (IETF) has worked out a Request For Comment (RFC) "Transmission of IPv6 via the IPv6 Convergence Sub-layer over IEEE 802.16 Networks" (RFC 5121, 2008) which provides a full conformant IPv6 connectivity through an IP point to point link. This solution fits the general business use case where each subscriber resides in its own sub-network. However, the requirements of the sub-network in an aeronautical environment might be different than the one of an ordinary business use case. Running IPv6 over AeroMACS shall be fully compliant to the IP standard, thereby, IP multicast shall be supported preferably in an efficient manner. It might also be desirable to support multicast at link layer which is difficult with point to point links. The problem statement and possible different solutions are discussed in the following.

First of all it is important to identify the relevant concepts of IPv6 addressing. In IPv6, nodes are attached to an access network via an interface, which is given at least one IPv6 address (i.e. the link local unicast address). Within this context a node can be understood as a device which implements IPv6. This means that an interface gets one or more IPv6 addresses assigned and not the node itself, which is a fundamental concept of IP. In other words a node may host several network interfaces which have different addresses. Thereby, the same node may be reachable through different IPv6 addresses.

An IPv6 capable node must be able of configuring its IPv6 address autonomously. An IPv6 address is created through a valid interface identifier and a valid subnet prefix. The subnet prefix may be a constant link local prefix (i.e. FE80::0), an advertised prefix received by *Router Advertisements,* or a prefix by a DHCPv6 server. The prefix is only valid on the link on which it is received - the prefix shall not be used on different links. Link local addresses allow communications between devices on a local link; such addresses cannot be used to communicate outside a network link.

Native multicast capability can be described through the following general concepts of the IP addressing model - first through the concept of a link and secondly through the concept of a subnet. A link is a term used to refer to a *topological area bounded by routers that decrement the IPv6 Hop Limit when forwarding a packet (c.f. RFC 4903, 2007).* The term subnet is *generally used to refer to a topological area that uses the same address prefix, where that prefix is not further subdivided except into individual addresses (c.f. RFC 4903, 2007).* Thereby, it is important to

recognize that *IPv6 continues the IPv4 model that a subnet is associated with one link. Multiple subnets may be assigned to the same link (c.f. RFC 4291, 2006).* Ideally, the Data Link layer addressing mechanisms can be directly used for the Internet Protocol addressing method. Some of the Internet layer protocols (e.g. Address Auto-configuration (RFC 4862, 2007), Neighbour Discovery (RFC 4861, 2007), Dynamic Host Configuration Protocol (RFC 315, 2003), or more generally protocols used for service discovery or name resolution) require native multicast capability of the underlying link, that is data packets can be distributed to all interested nodes on the same link without a decrement of the IPv6 Hop Limit field. If such a native multicast capability is not given by a certain link technology, an IP link model has to be presented towards the Internet layer which fulfils this requirement.

In principle, if a physical link characteristic is problematic at the Internet layer, mechanisms have to be defined that the link model appears properly at the Internet layer. The Internet Architecture Board (IAB) recommends using one of the two following models: The multi-access link model or the point to point link model. These models, if implemented properly, have no problems regarding the IP addressing model and the native multicast capability (c.f. RFC 4903, 2007).

3.1 IP point-to-point link model

Figure 12 considers a generic configuration of the IP point to point link model. Here exactly two nodes (i.e. Node A and Node B) are located on the same link. Native multicast is supported, that is, both nodes on the link are able to receive data packets which are sent to a link local multicast address. Also, both nodes can communicate with each other without any IPv6 Hop Limit decrement. Any Layer 2 device (e.g. bridges, switches, etc.) connected in the middle of the link is allowed. In terms of AeroMACS Node A would reflect the Mobile Station (MS) and Node B would be an Access Router, respectively.
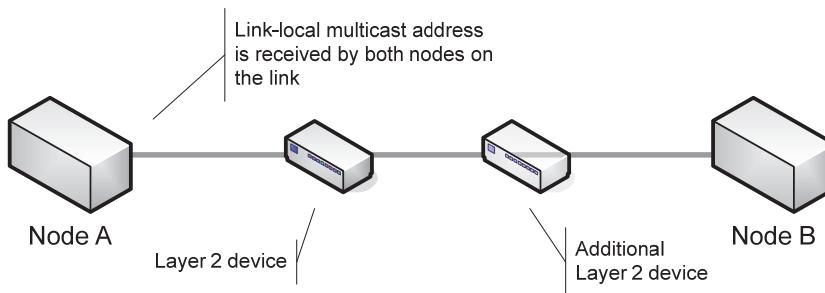
Fig. 12. A generic configuration of a "Point-to-Point Link Model".

3.1.1 Single prefix per mobile station

Considering the IP point to point link model presented to the network layer one possibility is to provide a single prefix to each Mobile Station. The AeroMACS communication system may make use of the IP convergence sub-layer (CS) or the Ethernet convergence sub-layer. In principle both solutions would work for the point to point connection, however, from a standard's point of view the two components are distinguished as the Ethernet configuration would allow a bridging functionality which the IP configuration does not.

Examining the case where data traffic is transmitted from the Mobile Station to the access router the following steps occur (cf. Figure 13):

1. An IPv6 packet is handed over to the convergence sub-layer. There the protocol fields of the higher layer data packet map to a Service Flow Identifier (SFID) that further maps to a Connection Identifier (CID). The connection identified through the CID is used to transmit the higher layer data packet to the Base Station (BS) - which is a layer two device as depicted in the Figure 12.
2. The "Data Path Function" of the BS encapsulates the IPv6 packet into a Generic Routing Encapsulation (GRE) tunnel. A unique GRE tag maps to the SFID of the connection.
3. The "Data Path Function" of the ASN Gateway receives the IPv6 packet and forwards it accordingly.

The other direction works analog. The "Data Path Function" of the ASN GW and the BS use the unique GRE tag to map to the SFID and CID of a connection, respectively. In such a way a virtual tunnel from the Access Router to the Mobile Station and vice versa is created. The "Data Path Function" creates the tunnel on a per service flow granularity. From an IP point of view these GRE tunnels have to be presented to the IP layer as "virtual interfaces" as each interface is more or less a single link. In such a way distinct prefixes have to be advertised on each link making the link a pure point to point link at layer 2 and layer 3. The main drawback of that solution is that standard multicast capabilities of the AeroMACS sub-network are disabled by default.
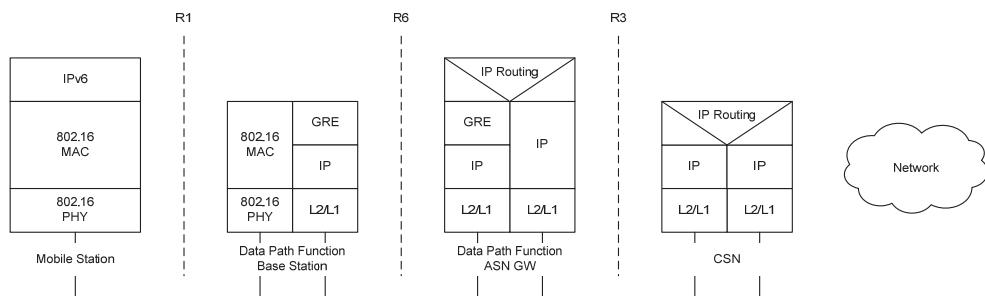

Fig. 13. Protocol stacks of different entities, considering an IP point to point link model.

3.1.2 Shared prefix for all mobile stations

If a shared prefix is used for all "Virtual Interfaces" the IP link model is violated, as the same prefix is advertised on different links. If only a single IP prefix is used (and native multicast among all attached Mobile Stations is being supported) an additional function is required (some kind of backend process). If a shared prefix is being needed for an AeroMACS system, the AeroMACS link as such has to be presented as a single link to the IP process. Therefore, some kind of backend process (or a similar approach like MLD snooping) is required. Such a backend process is no standard solution and would require a separate specification and implementation - such a solution is probably not desired.

3.2 IP multi-access link model

Figure 14 shows a generic configuration of a multi-access link model. One link is shared by one or more nodes (i.e. Node A, Node B, Node X, etc.). Thereby, the link may be attached to one or more routers. However, a router is not stringently necessary. Native multicast is supported, that is, all interested nodes on the link are able to receive data packets which are sent to a link local multicast address. Additionally, two nodes on the same link can

communicate with each other without any IPv6 Hop Limit decrement. Any Layer 2 device (e.g. bridges, switches, etc.) connected in the middle of the link is allowed.

The IEEE 802.16 over Ethernet option would provide such functionality - realized through the Ethernet CS interface and an Ethernet bridge (realized as layer 2 device). The AeroMACS link is presented to the IPv6 capable router only as a single interface. Note that the presentation of the single interface to the IP process as such has to be handled by the "Data Path Function" similar to the "backend process" presented earlier. The "Data Path Function" will encapsulate the data packets into a GRE tunnel and map them accordingly to a SFID and CID, respectively. Such a solution would be only possible if the Ethernet CS option is being used as the IP CS solution does not offer any possibility to bridge the data traffic.
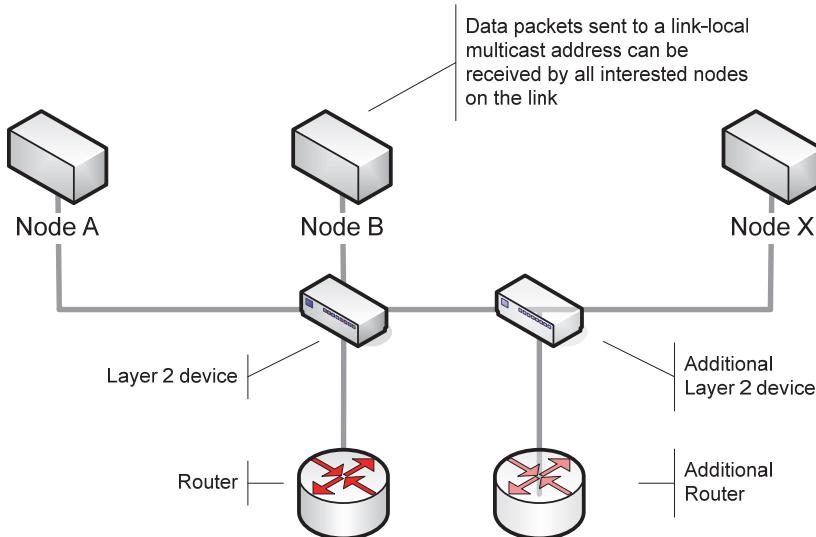
Fig. 14. A generic configuration of a "Multi-Access Link Model".

Alternatively the general encapsulation convergence sub-layer could be used to come up with a specific solution for the aviation case. However, this is no standard solution and would require additional development resources. A process which is not desired.

3.3 IP multicast

If one of the above mentioned IP models is implemented properly for the AeroMACS communication system IP multicast is always possible, as the system is fully IPv6 compliant. However, the advantage of IP multicast should be to transfer the data packet only as often

as necessary. Utilizing an IP point to point link model means that each multicast data packet transmitted to a group of multicast listeners belonging to the same cell and / or sector of a Base Station needs to be multiplied by the amount of listeners. Utilizing the Multi-Access link model means that the data packet is transmitted only once to the same group of multicast listeners belonging to the same cell/sector of a Base Station - this might increase the efficiency of an AeroMACS communication system.

In standard business cases of the mobile communication industry this may have no large impact. However, the aeronautical business case may have a larger interest in multicast applications. Therefore, the advantages and drawbacks of the different solutions shall be analyzed in depth in the future.

4. Future data traffic

Different types of application may utilize the AeroMACS communication system, among these could be fixed users, mobile users, and aircraft. By the term fixed users airport LAN extension could be considered such as unique equipment (terminals, cameras, etc.). Also Aeronautical Navigation Service Provider (ANSP) managed equipment like area navigation (RNAV) systems could be interpreted as fixed user. Mobile users would be airport surface vehicles such as airport trucks (catering, maintenance, refueling trucks, etc.) or mobile terminals which support for instance voice over IP. Aircraft would utilize an AeroMACS system by applications such as Air Traffic Control (ATC), Aeronautical Operational Communications (AOC), or Aeronautical Administrative Communications (AAC) which may have a direct operational and safety impact.

In order to assess the performance of an AeroMACS communication system properly an assumption on the data traffic load for airport surface communications was necessary. This sub-chapter summarizes the results of the data traffic load analysis for aircraft and mobile users conducted during the course of the SANDRA project (SANDRA). Services as anticipated by the COCRv2 report (COCR, 2007) and the AOC data link dimensioning report (AOC, 2010) were considered. In order to simulate a data traffic model the following inputs were necessary:

- An air traffic model: A simulation of the aircraft movement on the airport. That is, departing aircraft moving from ramp to runway, arriving aircraft moving from runway to ramp, and ground vehicles.
- Supported applications: Data link applications envisaged for the airport domain and their trigger events. Note that most aeronautical data link applications are triggered by events related to the progress of the flight. The events are provided by the air traffic model.
- Scenarios: These are the different use cases of the airport data link. This relates mostly to the amount of aircraft serviced, and the various sets of supported applications.

The output of the data traffic model is the statistical description of the expected offered load at application layer (ISO/OSI layer 7). The offered load is the amount of data produced by the applications – this does not include any overhead of the transport layer, network layer, or data link layer. Details about the implemented model can be found in (Ehammer, 2011).

The outcome of the evaluation showed that ATC applications contribute insignificantly to the offered load (in the order of a couple kbits/second), as these applications are mostly short. However, certain AOC applications may contribute significantly to the overall load, especially those related to software updates or post flight procedures. Results showed that

an average offered load of several megabits per second is possible. Thus justifying an introduction of a broadband wireless communication system at airports.

5. Simulation results

Within the SANDRA project MAC performance simulations have been conducted in 2011. Initial results are presented within this sub-chapter. In principle there are two different sets of WiMAX functions defined by the Mobile System Profile (as of IEEE 802.16-2009) - i.e. a set of functions to establish point to point connections over which data can be exchanged and a set of functions to support mobility. Within this sub-chapter a performance evaluation regarding data exchange functions is given.

In order to perform the evaluation a tool chain as depicted below has been used. Thereby, the parameter set is defined through simulation scenarios. The simulation itself is built according to requirements defined by simulation scenarios. The statistic program uses the XML based output data produced by the simulation as input to generate HTML reports which shall contain all necessary statistics in order to assess the different functions of the AeroMACS MAC layer. A similar approach has been conducted in several projects using the method described in (Ehammer, 2008).
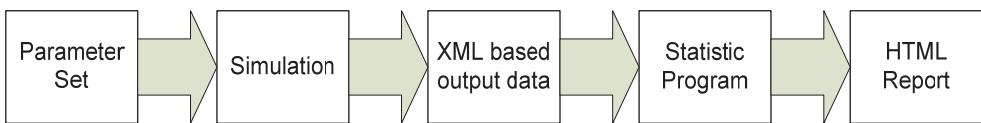
Fig. 15. Simulation tool chain.

The simulation effort concentrated solely on the MAC layer as depicted in Figure 16 below. The data transported through the connection (Mobile Station to Base Station) has been elaborated in a separate task (c.f. Chapter 3). However, the data traffic used for this evaluation is a synthetic data stream using a constant bit rate. The Physical Layer is modeled very simple through a uniformly distributed bit error rate (BER).
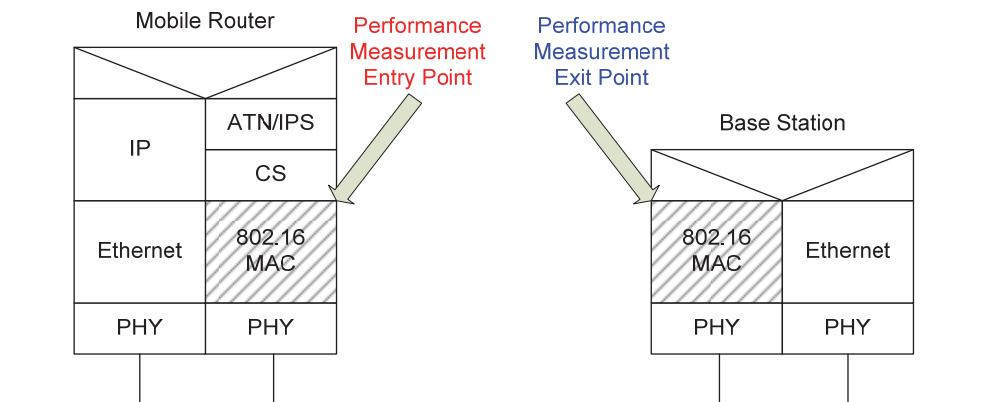
Fig. 16. Protocol stack of involved entities.

5.1 Evaluation

A basic set of parameter settings is evaluated against

- A varying bit error rate (the BER scenario).
- A varying load (the LOAD scenario)
- A varying amount of active users within the cell (the PIAC scenario).

Thereby, parameters such as latency, throughput, or loss are measured. Detailed evaluations show overhead figures resulted from the protocol itself and overhead caused through re-transmissions due to bad channel conditions. The evaluation scenario of varying bit error rates assumes a constant load and a constant amount of active users. Similarly, the evaluation scenario of varying load assumes a constant bit error rate and a constant amount of active users. Finally, the evaluation scenario of varying active users assumes a constant bit error rate and a constant load. For each evaluation a scenario is simulated several times in order to gain appropriate confidence intervals.

5.1.1 Reference scenario

The reference simulation scenario demonstrates basic data exchange capabilities of the AeroMACS protocol. The basic set of options necessary are the fragmentation and re-assembly options, the ARQ implementation with all allowed acknowledgment types set, a dynamic service flow addition capability (initiated by the Base Station after network entry), and the basic resource request and resource grant options.

The basic idea is to establish a data connection which has a QoS of "Best Effort". Additionally, this data connection shall support ARQ. Almost every time an aircraft has to transmit a data packet a resource request has to be issued - this is realized through the transmission of a CDMA ranging code via the ranging slot dedicated for periodic or bandwidth request ranging. If an aircraft has resources available it can also transmit a bandwidth request piggybacked via the "Grant Management Sub-header". Depending on the bit error rate, there will be loss. The sole bandwidth request header carries always aggregated bandwidth requests while the piggybacked bandwidth request carries incremented bandwidth requests. The maximum allowed PDU size is critical, especially if the bit error rate is high. It has to be coordinated with the ARQ block size as well. Obviously, larger ARQ block sizes than maximum fragment sizes do not make much sense.

The evaluation is based on a generic traffic generation where higher layer data packets have a typical IPv6 MTU size (i.e. 1500 bytes). The load per aircraft for uplink direction and downlink direction is assumed to be constant (evaluation in discrete steps from low load to high load). The amount of active participants (i.e. Mobile Stations or surface vehicles) per cell is assumed to be constant (evaluation in discrete steps from low to high). The bit error rate is assumed to be constant (evaluation in discrete steps, from good channel to bad channel conditions).

5.1.2 Simulation parameter settings

The simulation itself has a series of parameters to be considered the most important parameters for the BER evaluation scenario are briefly discussed. For the BER evaluation scenario 20 Mobile Stations are considered. Thereby, each MS shall generate on average a load of 60 kbps for the downlink direction and a load of 30 kbps for the uplink direction. The QoS parameters are set to Best Effort (BE). That means bandwidth request ranging opportunities have to be utilized if no resources are available when higher layer data

packets arrive. The maximum fragment size is set to 612 bytes. The ARQ settings are 128 bytes ARQ block size, 500 ms ARQ retry timeout, 10 seconds ARQ block lifetime, and all allowed acknowledgment types are enabled. Furthermore, acknowledgments may be piggybacked, bandwidth resource requests may be issued via the GSMH header (piggybacked), the compressed map feature is enabled, and packing of different MAC SDU is allowed. The BER values vary from 10^{-7} to 10^{-3}. The modulation and coding scheme has been chosen to be QPSK and CC rate 1/2. Different coding and modulation schemes should produce similar results except with higher throughput rates. Figure 17 shows an AeroMACS frame how it has been utilized for the presented simulation results. The DL-Prefix is present in each DL sub-frame as well as the DL and UL map. Downlink and Uplink Channel Descriptors (DCD/UCD) are transmitted periodically. Ranging opportunities such as initial and handover ranging slots or bandwidth request and periodic ranging slots are offered periodically as well. Dependent on the amount of BR ranging codes issued a variable amount of CDMA allocations have to be made available for Mobile Stations in the UL direction.
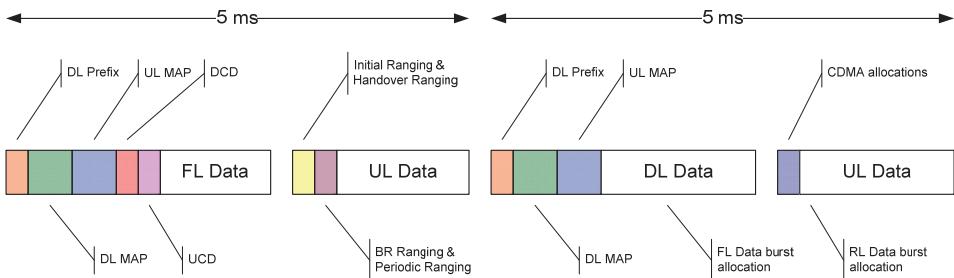

Fig. 17. Typical AeroMACS frame structure for this evaluation scenario.

5.1.3 Selected results

The figure below illustrate the higher layer (HiL) goodput and the data link layer (DLL) goodput. Goodput can be interpreted as user throughput, i.e. the number of useful bits per unit of time successfully forwarded by the network from a certain source address to a certain destination, excluding protocol overhead, and excluding retransmitted data packets. The figures underneath show the Forward Link (FL) and Reverse Link (RL) goodput. FL and RL are aeronautical terms where FL is the equivalent of DL and RL of UL, respectively. Furthermore, the figures underneath also show the FL and RL average latency for higher layer data packets and data link layer data packets. Note that the latency of data link layer packets is negligible as the transfer of a single MAC PDU takes at most 2 ms.

The FL results (i.e. Base Station to Mobile Station) show a controlled behavior until a BER of 10^{-5}. The higher layer data packets are still delivered at a BER of 5×10^{-5}, however the DLLgoodput increases due to re-transmissions caused by ARQ timeouts. As a result the average latency increases from approximately 100 ms to 1000 ms. Further decreasing the the channel (i.e. BER equals 10^{-4}) results in massive loss of higher layer data packets. packets. Higher layer packets are dropped as soon as the ARQ block lifetime of a MAC SDU (i.e. the higher layer data packet) expires. This is also reflected in the average FL latency which is

slightly underneath the ARQ block lifetime of 10 seconds (the graph accounts only for higher data layer packets which were delivered successfully). Decreasing the channel quality another time (i.e. BER equals 5×10^{-4}) results in total loss of data as well as almost no throughput of DLL data.

Considering the RL (i.e. Mobile Station to Base Station) a controlled behavior is only available until a BER of 10^{-5}. After that a similar behavior than the one of the FL can be observed. Due to the nature of the scheduling strategy used for the RL (i.e. best effort), acquiring new resources may be more time demanding, therefore an ARQ block lifetime timeout is more likely with a similar error rate than in the FL.
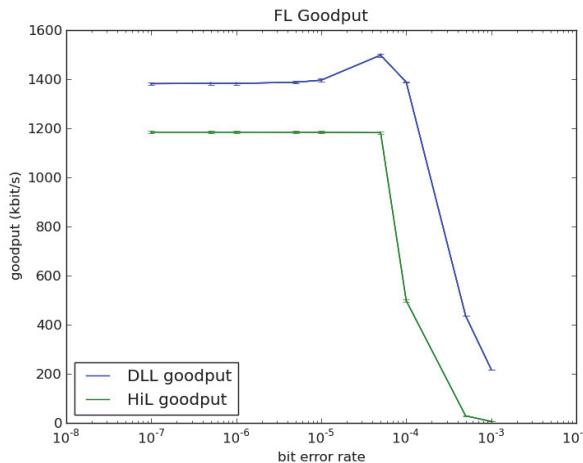
Fig. 18. FL Goodput - scenario BER.
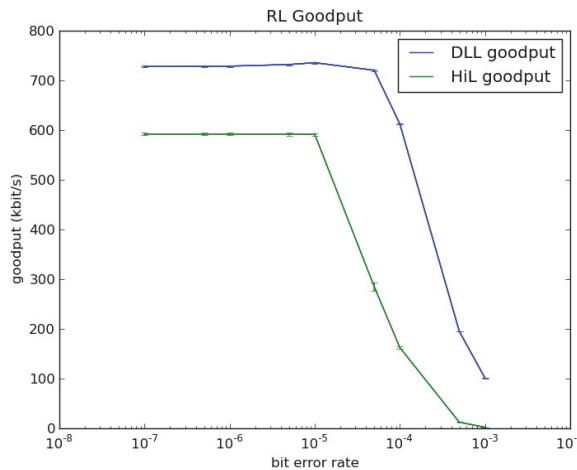
Fig. 19. RL Goodput - scenario BER.
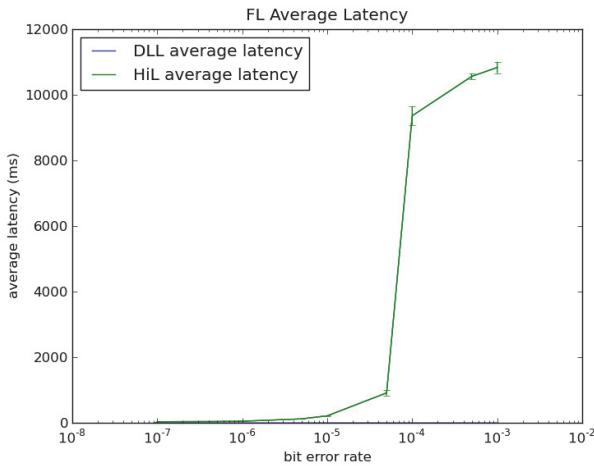
Fig. 20. FL avg. latency - scenario BER.
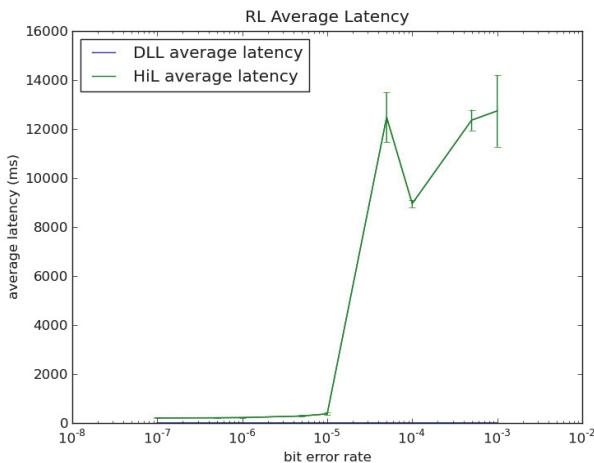
Fig. 21. RL avg. latency - scenario BER.

The LOAD evaluation scenario utilizes exactly the same parameter set than the BER evaluation scenario, except the bit error rate and the data traffic load. The LOAD scenario simulates a rather good channel condition, i.e. BER equal to 10^{-6}. The purpose of the LOAD scenario is to increase the load in steps starting from 9 kbps per Mobile Station RL data traffic and 18 kbps per Mobile Station FL data traffic. The intervals were set to 2 kbps for the RL and 4 kbps for the FL. Most load is generated by when 35 kbps for the RL and 70 kbps for the FL are set.

On the FL the higher layer data packets get all through, so do the data link layer packets. The DLL overhead stays proportional for all simulation scenarios. The average FL latency

increases slightly with more load but is still insignificantly. The RL goodput figure shows an interesting behaviour. Namely the slightly larger DLL goodput with low load. The reason therefore is that with low load resource request are less likely to be piggybacked. Thus, every time a higher layer data packet arrives a separate bandwidth request has to be issued. The overhead is significantly for lower loads as periodic ranging opportunities are accounted for DLL goodput. This overhead stays similar with higher loads, however, relatively the overall overhead decreases. The RL average latency starts to increase when the RL load starts to reach the capacity limit.
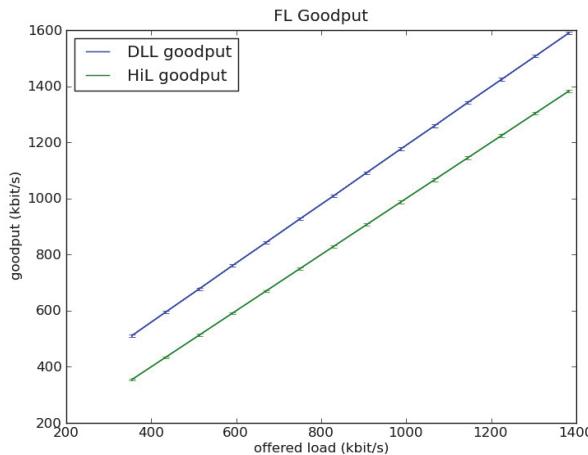

Fig. 22. FL goodput - scenario LOAD.
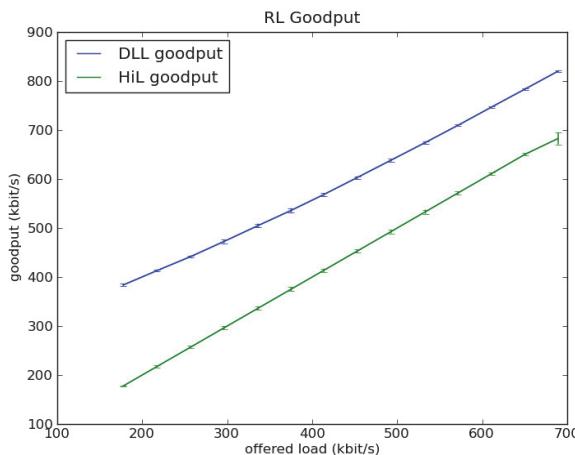
Fig. 23. RL goodput - scenario LOAD.
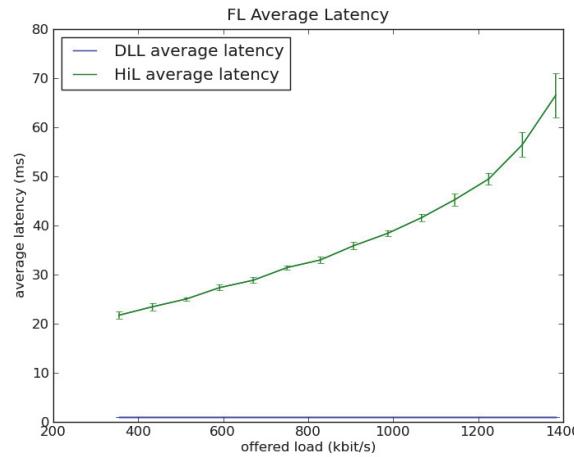
Fig. 24. FL avg. latency - scenario LOAD.
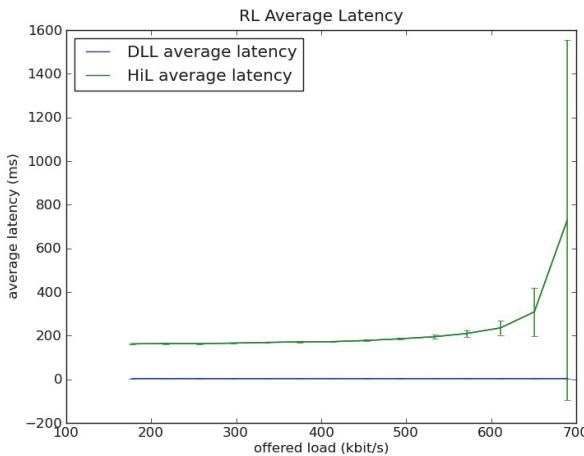
Fig. 25. RL avg. latency - scenario LOAD.

The PIAC (Peak Instantaneous Aircraft Count) evaluation scenario utilizes exactly the same parameter set than the LOAD evaluation scenario, except the amount of Mobile Stations and the data traffic load. The PIAC scenario simulates an average load of 20 kbps per Mobile Station FL data traffic and an average load of 10 kbps per Mobile Station RL data traffic. The amount of Mobile Stations is set to 10 and increased by intervals of 5 up to a maximum value of 80.

The FL higher layer goodput is reasonable until an amount of 70 concurrent users (and data traffic service flows). After that the throughput decreases and average latency figures increase. The RL higher layer goodput is stable until an amount of 50 concurrent users. However, the data link layer goodput increases with each interval. The reason therefore is

that bandwidth request ranging opportunities are served with unicast polls. The more concurrent Mobile Stations reside in a single cell, the more overhead is caused through this procedure. The higher layer goodput decreases after more than 50 Mobile Stations are trying to transmit concurrently. The DLL goodput is not reaching its maximum as the RL data sub-frame gets seriously fragmented through the unicast polls issued by the Base Station. As a result the scheduler gets serious problems utilizing the full spectrum sufficiently. Recall, that ARQ blocks have a fixed size. The average RL latency reflects the ARQ block lifetime.
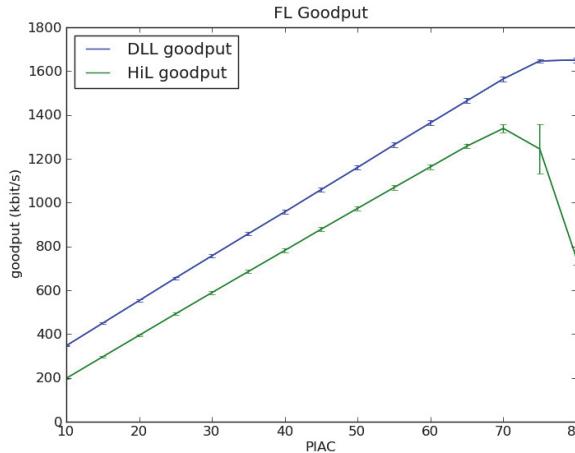

Fig. 26. FL goodput - scenario PIAC.
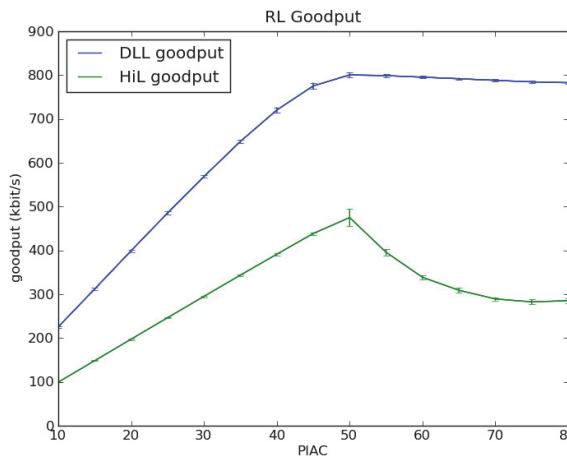
Fig. 27. RL goodput - scenario PIAC.
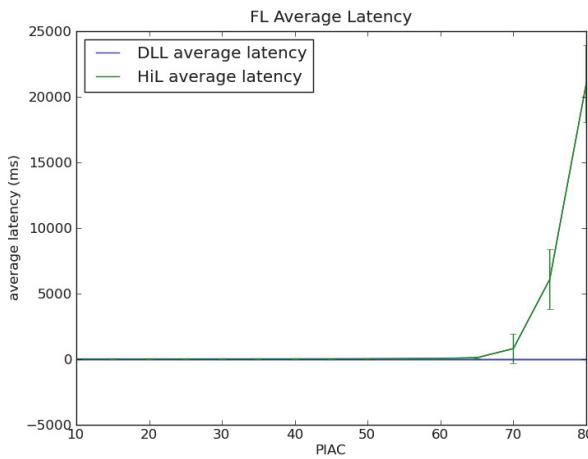
Fig. 28. FL avg. latency - scenario PIAC.
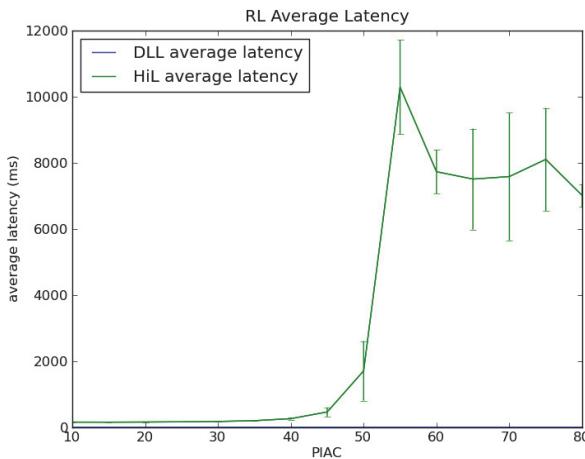
Fig. 29. RL avg. latency - scenario PIAC.

6. Conclusion

The demand to integrate the aircraft into the network centric concepts requires capable air ground data-links. AeroMACS shall provide this functionality at the airport surface. Infrastructure and equipage used for aeronautical procedures is evolving very slowly due to several reasons. Cost, interoperability, and safety issues are some among many reasons. Any new system integrated into the aeronautical environment will last for decades until it might eventually be replaced through a new system. Therefore, it is of importance to design new

systems carefully and with mature concepts in order to remain prepared for changing requirements in the future.

Integrating the AeroMACS sub-network into an IPv6 based aeronautical telecommunication network (ATN) is generally a problem which needs to be resolved from the application's point of view and from the operator point of view. It is also important to keep flexibility in order to be capable to adapt to any future changes of requirements. Especially if products have such long life cycles as in the aeronautical world it is almost impossible to assess the proper requirements.

Multicast applications may be very attractive to the future ATM concept, however, most of these applications are not realized yet (i.e. they exist only in theory). With a wrong sub-net configuration the introduction of application layer concepts based on multicast may be quite difficult and/or expensive.

During the course of the SANDRA project a data traffic load analysis has been conducted which showed that applications with significant load requirements would justify the introduction of a broadband wireless communication system for airport surface communications. Furthermore, MAC performance simulations have shown performance figures to be expected by a future AeroMACS system.

The current status of the AeroMACS profile is a draft. This means that further assessments on the maturity and performance of the technology shall clarify the suitability of AeroMACS for supporting the needs of future ATM concepts. Although currently prototypes are being implemented it is not believed that AeroMACS would be introduced before 2020. A realistic target for the deployment of an AeroMACS system is rather 2025 and beyond.

7. Acknowledgment

8. References

AOC (2010), M. Wood, S. Lebourg, B. Syren, and P. Huisman, SJU - AOC data-link dimensioning, edition 0.1, 2010.

COCR (2007), Communication Operating Concept and Requirements for the Future Radio System Version 2, May 2007.

Ehammer (2008). M. Ehammer, T. Gräupl, and C.H. Rokitansky, Applying SOA Concepts to the Simulation of Aeronautical Wireless Communication, Spring Simulation Multi-Conference 2008, April 2008.

Ehammer (2011). M. Ehammer, T. Gräupl, and E. Polo, AeroMACS Data Traffic Model, ICNS 2011, May 2011.

IEEE (2009). IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, IEEE Standards Association, May 2011, Available from <http://standards.ieee.org/about/get/802/802.16.html>.

RFC 3315 (2003). Droms et.al, RFC 3315 Dynamic Host Configuration Protocol for IPv6 (DHCPv6), July 2003, status: Proposed Standard, available at

< http://tools.ietf.org/html/rfc3315>.

RFC 4291 (2006). Hinden et al., RFC 4291 IP Version 6 Addressing Architecture, February 2006, status: Draft Standard, available at <http://tools.ietf.org/html/rfc4291>

RFC 4861 (2007). T. Narten et al., RFC 4861 Neighbor Discovery for IP version 6 (IPv6), September 2007, status: Draft Standard, available at <http://tools.ietf.org/html/rfc4861>.

RFC 4862 (2007). Thomson et al., RFC 4862 IPv6 Stateless Address Auto-configuration, September 2007, status: Draft Standard, available at <http://tools.ietf.org/html/rfc4862>.

RFC 4903 (2007). D. Thaler, RFC 4903 Multi-Link Subnet Issues, May 2007, status: Informational available at <http://tools.ietf.org/html/rfc4903>.

RFC 5121 (2008). Patil et al., RFC 5121 Transmission of IPv6 via the IPv6 Convergence Sub-layer over IEEE 802.16 Networks, February 2008, status: Proposed Standard, available at <http://tools.ietf.org/html/rfc5121>.

SANDRA. Seamless Aeronautical Networking through integration of Data links, Radios, and Antennas (SANDRA). Large Scale Integrated Project within FP7 - Grant Agreement n° 233679, accessible at <http://www.sandra.aero>.

Sayenko, A.; Tykhomyrov, V.; Martikainen, H. and Alanen, O. (2007). Performance analysis of the ieee 802.16 arq mechanism, *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, ISBN: 978-1-59593-851-0, New York, NY, USA.

WiMAX Forum (2009). WiMAX Forum Mobile System Profile Specification: Release 1.5 TDD Specific Part, August 2009.

WiMAX Forum (2010). WiMAX Forum Mobile System Profile DRAFT-T23-R010v09-B; Working Group Approved Revision, May 2010.

The LDACS1 Link Layer Design

Thomas Gräupl and Max Ehammer
University of Salzburg
Austria

1. Introduction

Air transportation is an important factor for the economic growth of the European Union, however, the current system is already approaching its capacity limits and needs to be reformed to meet the demands of further sustainable development (Commission of the European Communities, 2001). These limitations stem mainly from the current European air traffic control system.

Air traffic control within Europe is fragmented due to political frontiers into regions with different legal, operational, and regulative contexts. This fragmentation decreases the overall capacity of the European air traffic control system and, as the system is currently approaching its capacity limits, causes significant congestion of the airspace. According to the European Commission airspace congestion and the delays caused by it cost airlines between €1.3 and €1.9 billion a year (European Commission, 2011). For this reason, the European Commission agreed to adopt a set of measures on air traffic management to ensure the further growth and sustainable development of European air transportation.

The key enabler of this transformation is the establishment of a Single European Sky[1] (SES). The objective of the SES is to put an end to the fragmentation of the European airspace and to create an efficient and safe airspace without frontiers. This will be accomplished by merging national airspace regions into a single European Flight Information Region (FIR) within which air traffic services will be provided according to the same rules and procedures.

In addition to the fragmentation of the airspace the second limiting factor for the growth of European air transportation lies within the legacy Air Traffic Control (ATC) concept. In the current ATC system, which has been developed during the first half of the twentieth century, aircraft fly on fixed airways and change course only over navigation waypoints (e.g. radio beacons). This causes non-optimal paths as aircraft cannot fly directly to their destination and results in a considerable waste of fuel and time[2]. In addition, it concentrates aircraft onto airways requiring ATC controllers to ascertain their safe separation.

The tactical control of aircraft by ATC controllers generates a high demand of voice communication which is proportional to the amount of air traffic. As voice communication

---

[1]Regulation (EC) No 549/2004 of the European Parliament and of the Council of 10 March 2004.
[2]On average, flight routes within Europe are 49 kilometres too long (European Commission, 2011). EUROCONTROL reported 9,916,000 IFR (Instrument Flight Rules) flights in 2007 resulting in 485,884,000 unnecessary flight kilometres over Europe.

puts a considerable workload on the human controller the air traffic cannot be increased arbitrarily without compromising the safety of the system. This situation is made worse by the fact that the radio spectrum dedicated to aeronautical voice communication is becoming increasingly saturated i.e. even if the human controllers could cope with more air traffic safely, there would not be enough voice frequencies to do so. Excessive controller workload and voice frequency depletion are therefore the main technical problems of the current air traffic control system.

The introduction of advanced Air Traffic Management (ATM) procedures and automated support tools will significantly decrease the controller workload. However, advanced ATM requires aircraft to be equipped with accurate position determination and collision avoidance equipment as well as data communications to integrate them into the ATM, System Wide Information Management (SWIM) and Collaborative Decision Making (CDM) processes (Helfrick, 2007).

Data communications is required as ATM transfers parts of the decision making from air traffic controllers to cockpit crews supported by automated procedures and algorithms (e.g. self-separation). The aircrews must now be provided with timely, accurate, and sufficient data to gain the situational awareness necessary to effectively collaborate in the collaborative decision making process of ATM. This requires the availability of sufficiently capable data links. However, the data link solutions available today cannot provide the capacity and quality of service required for the envisaged system wide information management (Eleventh Air Navigation Conference, 2003). Improved air-ground communication has therefore been identified as one key enabler in the transformation of the current air transportation system to an ATM based Single European Sky.

2. Development of LDACS1

Today's air-ground communication system is based on analogue VHF voice transmission and is used for tactical aircraft guidance. It is supplemented by several types of aeronautical data links that are also operated in the VHF COM band, most notably ACARS (FANS 1/A) and VHF Digital Link Mode 2.

However, these data links are scarcely deployed. Their further deployment is blocked by the fact that the VHF band is already heavily used by voice communication and is anticipated to become increasingly saturated in high density areas (Kamali, 2010). Introducing additional communication systems into the same frequency band will therefore increase the pressure on the existing infrastructure even further. ACARS and VDL Mode 2 can therefore not provide a viable upgrade path to ATM.

At the eleventh ICAO Air Navigation Conference in 2003 it has therefore been agreed that the aeronautical air-ground communications infrastructure has to evolve in order to provide the capacity and quality of service required to support the evolving air traffic management requirements.

It was the position of the airlines (represented by IATA) that the "air-ground infrastructure should converge to a single globally harmonized, compatible and interoperable system" (IATA, 2003). Thus FAA and EUROCONTROL, representing the regions feeling the most pressure to reform their air-ground communication infrastructure, initiated the Action Plan (AP17) activity to jointly identify and assess candidates for future aeronautical communication systems (EUROCONTROL & FAA, 2007a). This activity was coordinated with the relevant stakeholders in the U.S. (Joint Planning and Development Office Next

Generation Air Transportation System; NextGen) and in Europe (Single European Sky ATM Research; SESAR).

Action Plan 17 concluded in November 2007 and comprised six technical tasks and three business tasks. The business tasks are not of relevance in the context of this chapter, however, the technical tasks were:

- Task 1: Improvements to current systems - frequency management
- Task 2: Identify the mobile communication operational concept
- Task 3: Investigate new technologies for mobile communication
- Task 4: Identify the communication roadmap
- Task 5: Investigate feasibility of airborne communication flexible architecture
- Task 6: Identify the Spectrum bands for new system

The data link technology discussed in this chapter (LDACS1) was developed as input to AP17 Task 3 and its follow-up activities (Gräupl et al., 2009). As one follow-up activity to AP17, EUROCONTROL funded the development and first specification of the LDACS1 system. Although there was no formal cooperation between EUROCONTROL and FAA at this point (AP17 had already been concluded) the development of LDACS1 was observed and advised by FAA and its sub-contractors NASA, ITT and the MITRE cooperation (Budinger et al., 2011).

After the end of the EUROCONTROL funded initial specification the development of the LDACS1 technology was continued in the "Consolidated LDACS1 based on B-AMC" CoLB project of the Austrian research promotion agency FFG as part of the TAKEOFF program. This project produced an updated specification and extensive guidance material. The overview paper (Kamali, 2010) provides an independent summary of the development of the L-DACS systems up to the year 2010. In 2011 the development of LDACS1 was continued in the framework of the SESAR Programme (Sajatovic et al., 2011).

2.1 Design goals

The primary design goals of the LDACS1 technology proposal were defined by the high level objectives formulated in AP17 (Fistas, 2009):

- The system development shall be facilitated and expedited through the choice of appropriate components and mature standards.
- The new system should be capable to operate in the L-band without interfering with existing users of the band.
- The system performance should meet the requirements defined in AP17 technical task 2.

The reason for the first design goal was the target deployment year of the future radio system, 2020. The aeronautical industry has comparatively long deployment cycles: In the past the deployment of safety related communication systems has taken between 8 to 15 years i.e. it is required that any future radio system candidate has already achieved a sufficiently high maturity by now, if its initial deployment shall begin by 2015. Starting deployment in 2015 shall allow for a period of pre-operational use before operational service starts in 2020.

Meeting the requirements defined in AP17 technical task 2 requires to support operational aeronautical communication i.e. Air Traffic Services (ATS) and Aeronautical Operational Control (AOC) communications. ATS communication provides navigation, control and situational awareness, while AOC communication is used to perform the business

operations of the airline. The system shall be capable to provide simultaneous ATS and AOC communication with adequate performance as of 2020 and beyond. Due to regulatory reasons passenger communication is out of scope of LDACS1.

These three high level objectives of AP17 were augmented by a number of non-technical, legal and political requirements, which are not discussed here. Within this chapter only the design aspects and evaluation criteria related to the performance of the system are discussed in detail. This was reflected in the identification of five relevant design goals.

Responsiveness is the capability of the system to react to communication demand in accordance with given requirements. This comprises the ability to deliver data traffic within specified delays and to provide swift voice service with minimum latency.

Reliability is the ability of the system to transmit data without losing or duplicating information. The required level of reliability is expressed in terms of service continuity.

Scalability is required for the future radio system in order to handle growing amounts of data traffic and users i.e. the technology should support as many use cases identified in AP17 technical task 2 as possible with acceptable quality of service.

Efficient resource usage of the new system is dictated by the scarcity of the available spectrum. This implies avoiding unnecessary protocol overhead (e.g. finding the right balance between forward error correction and backward error correction) and fair distribution of channel resources among users with the same priority.

Resilience is the ability of the future radio system to provide and maintain an acceptable quality of service even under adverse conditions. In particular this refers to periods of excessive load and high numbers of users. The system shall behave predictable and, if it fails, this must be detected early and reported immediately.

Of the five design goals presented above only the first three are discussed in detail in this chapter. The last two are touched only briefly. Note that the Communications Operating Concepts and Requirements (COCRv2) document (EUROCONTROL & FAA, 2007b), which was another output of AP17 technical task 2, defines validation criteria for one-way latency (TT95-1 way), continuity, integrity, and availability. These criteria define the target parameters of the L-DACS design and are related to the validation parameters discussed in section 4.3.

3. Design analysis

LDACS1 was designed to provide an air-ground data link with optional support for digital air-ground voice. It is optimized for data communication and designed to simultaneously support ATS and AOC communications services as defined in EUROCONTROL's and FAA's "Communication Operating Concept and Requirements for the Future Radio System" (EUROCONTROL & FAA, 2007b).

The key features of LDACS1 are:

- Cellular radio system with up to 512 users per cell. Up to 200 nautical miles range.
- Frequency division duplex with adaptive coding and modulation providing from 303.3 kbit/s up to 1,373.3 kbit/s in each direction.
- Acknowledged and unacknowledged point-to-point communication between ground-station and aircraft-station.
- Unacknowledged multicast communication between ground-station and aircraft-stations (ground-to-air direction only).

• Hierarchical sub-network architecture with transparent handovers between radio cells. This chapter discusses only the protocols of the wireless part of the LDACS1 system i.e. the air interface between the ground-station and the aircraft-station. Physical layer details, sub-network architecture, cell entry, and handovers are not discussed here.

3.1 Functional architecture

The LDACS1 air-ground communication architecture is a cellular point-to-multipoint system with a star-topology where aircraft-stations are connected to a ground-station via a full duplex radio link. The ground-station is the centralized instance controlling the air-ground communications within a certain volume of space called an LDACS1 cell. The LDACS1 protocol stack defines two layers, physical layer and data link layer (comprising two sub-layers itself) as illustrated in Fig. 1.
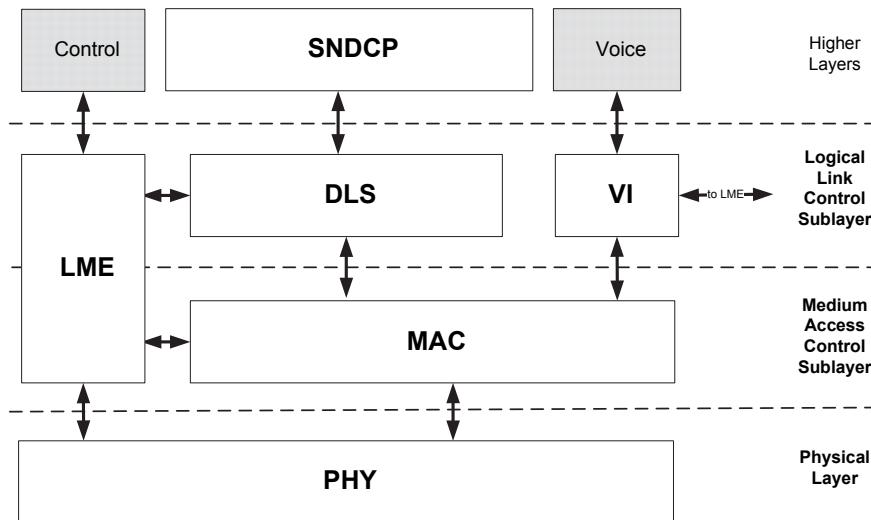
Fig. 1. LDACS1 protocol stack.

The physical layer provides the means to transfer data over the radio channel. The LDACS1 ground-station simultaneously supports bi-directional links to multiple aircraft-stations under its control. The forward link direction (FL; ground-to-air) and the reverse link direction (RL; air-to-ground) are separated by frequency division duplex (FDD). In the RL direction different aircraft-stations are separated in time (using time division multiple access; TDMA) and frequency (using orthogonal frequency division multiple access; OFDMA).

The ground-station transmits a continuously stream of OFDM symbols on the forward link. Aircraft-stations transmit discontinuous on the RL with radio bursts sent in precisely defined transmission opportunities using resources allocated by the ground-station. An aircraft-station accesses the RL channel autonomously only during cell-entry. All other reverse link transmissions, including control and user data, are scheduled and controlled by the ground-station.

The data-link layer provides the necessary protocols to facilitate concurrent and reliable data transfer for multiple users. The functional blocks of the LDACS1 data link layer architecture

are organized in two sub-layers: The medium access sub-layer and the logical link control sub-layer (LLC). The logical link control sub-layer manages the radio link and offers a bearer service with different classes of service to the higher layers. It comprises the Data Link Services (DLS), and the Voice Interface (VI). The medium access sub-layer contains only the Medium Access (MAC) entity. Cross-layer management is provided by the Link Management Entity (LME). The Sub-Network Dependent Convergence Protocol (SNDCP) provides the interface to the higher layers.

The MAC entity of the medium access sub-layer manages the access of the LLC entities to the resources of the physical layer. It provides the logical link control sub-layer with the ability to transmit user and control data over logical channels. The peer LLC entities communicate only over logical channels and have no concept of the underlying physical layer.

Prior to fully utilizing the system, an aircraft-station has to register at the controlling ground-station in order to get a statically assigned dedicated control channel for the exchange of control data with the ground-station. The ground-station dynamically allocates the resources for user data channels according to the current demand as signalled by the aircraft-stations.

Except for the initial cell-entry procedure all communication between the aircraft-stations and the controlling ground-station (including procedures for requesting and allocating resources for user data transmission and retransmission timer management), is fully deterministic and managed by the ground-station. Under constant load, the system performance depends only on the number of aircraft-stations serviced by the particular ground-station and linearly decreases with increasing number of aircraft.
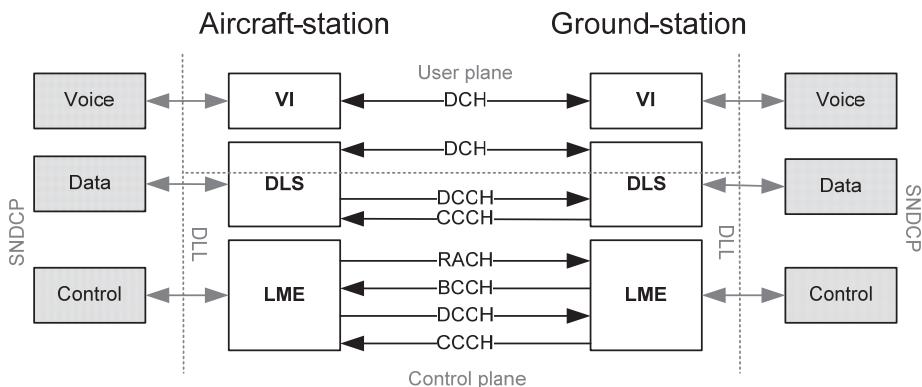
Fig. 2. L- DACS 1 logical channel structure.

Bidirectional exchange of user data between the ground-station and the aircraft-station is performed by the Data Link Service (DLS) entity using the logical data channel (DCH) for user plane transmissions[3]. Control plane transmissions from the aircraft-station to the ground-station are performed over the logical dedicated control channel (DCCH). Ground-to-air control information is transmitted in the common control channel. The random access

---

[3]Note that the Voice Interface (VI) also uses the DCH for its transmissions.

channel (RACH) and the broadcast control channel (BCCH) are used for cell-entry, cell-exit, and handover. The relation of the logical channels to the functional blocks of the LDACS1 logical link control layer is illustrated in Fig. 2.

The Data Link Service (DLS) provides the acknowledged and unacknowledged exchange of user data over the point-to-point reverse link or point-to-multipoint forward link. There is one DLS in the aircraft-station and one peer DLS for each aircraft-station in the ground-station.

The ground-station Link Management Entity (LME) provides centralized resource management for LDACS1. It assigns transmission resources, provides mobility management and link maintenance. It assigns forward link and reverse link resources taking channel occupancy limitations (e.g. limiting the aircraft-station duty cycle to minimize co-site interference) into account. In addition, the LME provides dynamic link maintenance services (power, frequency and time adjustments) and supports Adaptive Coding and Modulation (ACM).

The Voice Interface (VI) provides support for virtual voice circuits. The voice interface provides only the transmission and reception services, while LME performs creation and selection of voice circuits. Voice circuits may either be set-up permanently by the ground-station LME to emulate party-line voice or may be created on demand.

LDACS1 shall become a sub-network of the Aeronautical Telecommunications Network (ATN). The Subnetwork Dependent Convergence Protocol (SNDCP) provides the LDACS1 interface to the network layer and a network layer adaptation service required for transparent transfer of Network layer Protocol Data Units (N-PDUs) of possibly different network protocols (ATN/IPS and ATN/OSI). The SNDCP should also provide compression and encryption services required for improving and securing the wireless channel.

3.2 Input from other systems

Most features of the LDACS1 data link layer design are based on the experience gained from the precursor system B-AMC (Rokitansky et al., 2007). The most important protocol element adopted from B-AMC is the medium access approach. The protocol stack architecture and the data link service protocol were redesigned on the basis of the lessons learnt from B-AMC.

However, a considerable amount of input was also received from other AP17 candidate systems. Probably the most influential external input[4] to the LDACS1 design came from the TIA-902 P34 standard. The message formats of the medium access layer and the addressing scheme were directly derived from this system (Haindl et al., 2009). The concept of OFDM tiles and FL and RL allocation maps was adopted from the WiMAX standard. Additional input from WiMAX has gone into the design of the physical layer.

3.3 Physical layer overview

LDACS1 is intended to operate in 500 KHz wide cannels located in the 1 MHz gaps between adjacent DME[5] channels in the L-band. This type of design is called an inlay system. Inlay systems and similar methods of utilizing "white"-space spectrum are an approach to frequency allocation receiving increased interest, as finding free ("green") spectrum

---

[4]Kindly supported by FAA, NASA, MITRE, and ITT.
[5]Distant Measuring Equipment (DME) is an aeronautical radio navigation system.

becomes progressively more difficult. LDACS1 shall cover the needs for aeronautical data communication well beyond the year 2030. Therefore it is necessary to make as much bandwidth as possible available to the system. As the L-band is already crowded by other aeronautical and military systems, an inlay concept not requiring any green spectrum is an attractive approach.[6]

However, designing and deploying an inlay technology is a non-trivial matter as co-existence with legacy systems has to be ensured. The problem of co-existence can be decomposed into two parts: Interference from the inlay system towards the legacy systems and interference from the legacy systems towards the inlay system.

Naturally the new system must not disturb the operation of the existing infrastructure. The legacy systems can, however, not be modified, thus, the inlay system has to carry most of this burden. LDACS1 uses a powerful combination of different methods for side-lobe suppression and reduction of out-of-band radiation described in (Brandes, 2009).

The second part of the problem is to design the inlay system robust against interference from existing systems. This is a non-trivial task as many deployed legacy systems have sub-optimal interference characteristics according to modern standards. Most inlay designs therefore try to mitigate the interference of the existing system using sophisticated signal processing and error correcting codes. This is also the approach taken by LDACS1.

The two parts of the co-existence problem cannot be seen in isolation. Any approach to one of both problems has consequences for the other. Therefore it is necessary to find an integrated solution. Depending on the efficiency of the mutual interference suppression two types of inlay systems are possible: The first type is an inlay system that can be deployed completely independent of existing systems. This is an ideal case that can seldom be achieved. The second type of inlay system requires a certain level of coordination.

Close inspection of L-band spectrum usage reveals that the range form 962 MHz to 1025 MHz and 1150 MHz to 1213 MHz is used only for DME reply channels i.e. only the DME transponder sites will use these frequencies for transmissions. Therefore it can be assumed that an LDACS1 ground-station transmitting in the same region will most likely not disturb a nearby (in terms of frequency and distance) airborne DME receiver. Consequently, as a first measure to reduce interference between both systems, LDACS1 was designed as a Frequency Division Duplex (FDD) system[7]. The LDACS1 forward link (FL; ground-to-air) is transmitted in the same region of spectrum as the DME reply (i.e. ground-to-air) channels, from 985 MHz to 1009 MHz. This respects safety margins for the universal access transceiver (UAT), secondary surveillance radar (SSR), and global navigation satellite systems.

Finding an appropriate spectrum allocation for the LDACS1 Reverse Link (RL; air-to-ground) is less obvious as there is no region exclusively in use by DME interrogation channels. Respecting safety margins for the critical systems, two candidate intervals remain: 1048 MHz to 1072 MHz and 1111 MHz to 1135 MHz. As the first option is currently less used by DME, the LDACS1 RL has been allocated in this region (1048 MHz to 1072 MHz). This allows for 24 L-DACS1 FDD channel pairs. The second region is considered as optional extension for now.

The LDACS1 OFDM parameters were chosen according to the characteristics of the aeronautical mobile L-band channel (Brandes, 2009). The forward link and reverse link

---

[6] Note, however, that L-DACS1 can also be deployed in green spectrum without any changes to the technology.

[7] Another reason for the use of FDD was to avoid the large guard interval required between the FL and RL section of TDD.

channels have an effective bandwidth of 498.05 kHz each. Within that bandwidth, 50 OFDM sub-carriers are placed, separated by 9.765625 kHz. Each sub-carrier is separately modulated with a symbol duration of T_s= 120 µs.

LDACS1 employs concatenated block coding and Reed-Solomon coding in the physical layer. Using the default coding and modulation[8] (QPSK, coding rate 0.45) LDACS1 provides a data rate of 303.3 kbit/s in each direction. Using more aggressive coding and modulation schemes this can be increase up to 1373.3 kbit/s (64QAM, coding rate 0.68) in each direction. Different aircraft-stations may employ different coding and modulation schemes using adaptive coding and modulation (ACM).

The physical layer design includes propagation guard times sufficient for a maximum range of 200 nautical miles. In real deployments the LDACS1 maximum transmission power may, however, have to be limited in order to protect receivers of other L-band systems.

3.3.1 Frame structure

The LDACS1 protocol structures the physical layer on the basis of OFDM frames. Frames are combined into multi-frames and super-frames. The LDACS1 super-frame is the highest element of the physical layer framing hierarchy. The super-frame duration is 240 milliseconds or 2000 OFDM symbols. This is a multiple of the voice sample length (20 milliseconds) produced by the AMBE ATC10B vocoder[9]. A forward link super-frame comprises a broadcast control frame BC sub-divided into the BC1, BC2, and BC3 sub-frames and four multi-frames. A reverse link super-frame comprises two random-access opportunities in the random access RA frame and four multi-frames. The super-frame layout is illustrated in Fig. 3 and Fig. 4.
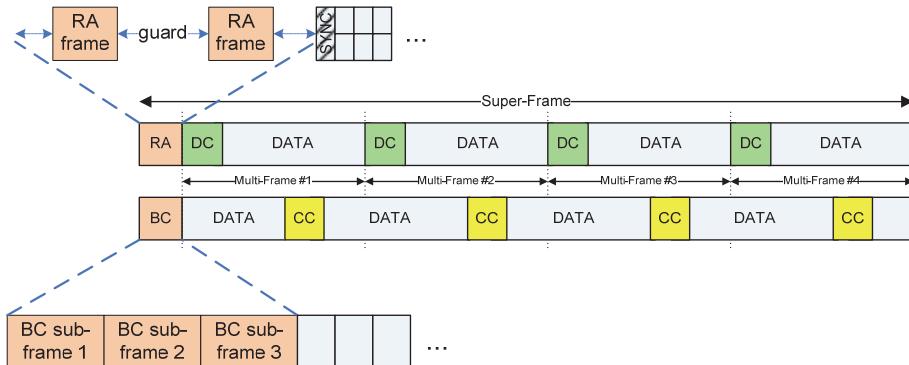
Fig. 3. LDACS1 super-frame structure.

The length of the forward link and reverse link multi-frames is 58.32 milliseconds or 486 symbols. Each forward link multi-frame comprises 9 Data OFDM frames. A variable number of these frames, starting with the fifth data frame, use a different channel coding and are designated as common control CC frames.

---

[8] The default coding and modulation was designed with the maximum DME sending rate (3600 pulse pairs per second) in mind.

[9] The AMBE ATC10B vocoder is the only digital vocoder currently certified for operational use in aeronautics.

Each RL Multi-Frame (MF) comprises one dedicated control DC slot and one Data slot. These slots are sub-divided into tiles. The reverse link DC slot starts with the RL synchronization symbols and the first two reverse link tiles. Its length is variable between two tiles and fifty-two tiles. The remaining RL tiles create the reverse link data slot.
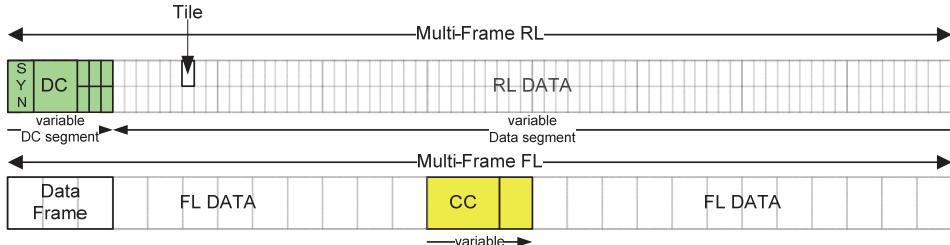
Fig. 4. LDACS1 multi-frame structure.

3.4 Medium access sub-layer

The MAC entity of the medium access sub-layer manages the access of the logical link control entities to the resources of the physical layer. The medium access sub-layer provides the logical link control sub-layer with the ability to transmit user and control data in logical channels.

3.4.1 Medium access

The medium access service supports the transmission of user and control data over logical channels. It manages the access of the logical link control sub-layer entities (DLS, LME, and VI) to the time slots conveying the logical channels. The broadcast control channel (BCCH) and the random access channel (RACH) are mapped to the BC and RA slots, respectively. The common control channel (CCCH) is conveyed in the CC slot of the forward link, the dedicated control channel (DCCH) in the DC slot of the reverse link. The forward and reverse data channels (DCH) are mapped to the corresponding data slots of the frame structure.

Since the forward link is exclusively used by the ground-station, no sophisticated multiple access scheme is required in this direction. The ground-station is the only user of the FL time slots. It can therefore allocate FL channel resources (i.e. bytes in the continuous FL transmission) locally according to the required quality of service. This allocation is announced to the aircraft-stations in the common control channel CCCH.

The RL uses a bandwidth on demand scheme. Aircraft-stations have to request channel resources (i.e. RL tiles) from the ground-station before they can transmit their data channel DCH in the RL Data slot. To this purpose aircraft-stations are polled by the ground-station for their resource request in round-robin. Up to 52 aircraft can be polled in one multi-frame. This approach makes resource requests deterministic and contention free. This procedure is illustrated in Fig. 5. Note that the size of the DC slot is variable. It can therefore be increased to provide the capacity necessary to transmit all resource requests.

The resource request is an aggregate request for all resources needed by the aircraft-station. The ground-station assigns resources according to the required quality of service using an appropriate scheduling algorithm. It has to keep track of allocations to avoid duplicate assignments. The coding and modulation of tiles in the RL Data slot can either be fixed for

the entire LDACS1 cell or be changed dynamically by the ground-station LME. The concept of the FL and RL resource allocation is illustrated in Fig. 6. Note that the size of the CC slot is variable. It can therefore be increased to provide the capacity necessary to transmit all resource allocations.
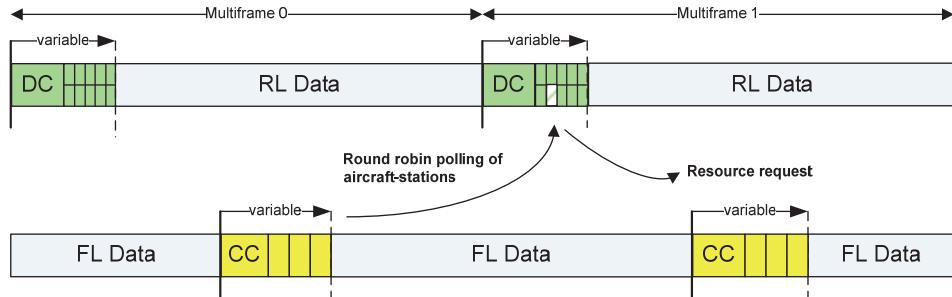
Fig. 5. RL resource request over the DCCH.
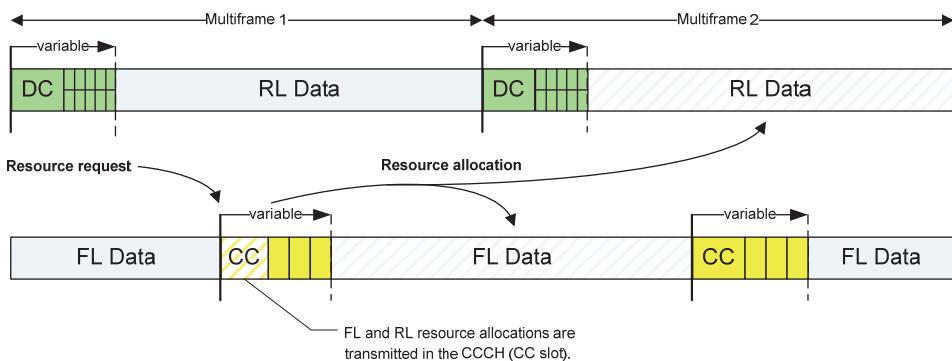
Fig. 6. RL resource allocation over the CCCH.

The LDACS1 medium access sub-layer does neither generate nor process the control messages (resource requests, resource allocations, etc.) transferred over the RACH, BCCH, DCCH and CCCH itself. Resource requests, resource allocations, acknowledgements, etc. are generated and consumed only by the LLC sub-layer. The format of these control messages is specific for each logical channel and documented in the LDACS1 specification (Sajatovic et al., 2011). The only exceptions from this rule are the FL and RL resource allocations of the CCCH, which are stored in the MAC. This is done to ensure that outgoing data is transmitted correctly (i.e. with the assigned coding and modulation) and to determine the source of incoming transmissions from the allocations.

3.4.2 Medium access analysis

The size of the CC and DC slots are variable. The optimal CC slot length is the smallest number of CC OFDM frames sufficient to convey all forward link control messages. Determining the optimal DC slot size is less trivial, however, the minimum DC slot size can be derived from the 95% percentile latency requirements (see Section 4.3) according to the

following design approach: If the physical layer is configured to provide a DLS packet error rate of less than five percent, 95% of the DLS packets can be delivered without a retransmission. The 95% percentile of the higher layer latency is then (approximately) equal to the MAC latency in this case i.e. if the MAC latency can be given as a function of the DC slot length, the minimum DC slot length can be derived from the 95% latency requirements[10] given in section 4.3.

The duration of a RL transmission TX_{RL} is bounded by

$$TX_{RL} \leq TX_{MAC} + TX_{Transmission} \tag{1}$$

where T_{MAC} is the medium access latency (resource request + resource allocation), and $T_{Transmission}$ is the transmission time of the data itself.
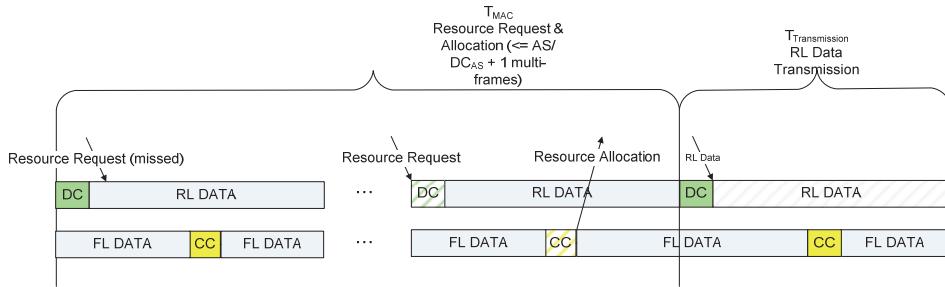
Fig. 7. Maximum length of RL packet transmission.

An aircraft-station can send a resource request in the DC slot only when it is polled. In the worst case, if the aircraft-station just missed this time slot, it has to wait for the current multi-frame and one complete reservation cycle (i.e. the time until it is polled again in round robin) to make a request. This is illustrated in Fig. 7. The corresponding resource allocation should[11] be transmitted in the next CC slot, which is within the last multi-frame of the reservation cycle. Thus T_{MAC} is bounded by the length of the reservation cycle.

$$T_{MAC} \leq \left\lceil \frac{AS}{DC_{AS}} \cdot MF \right\rceil + MF \tag{2}$$

where AS is the number of registered aircraft stations, DC_{AS} is the number of aircraft stations polled per DC, and MF is the average length of the multi-frame (60 milliseconds neglecting the RA/BC slot).

It is assumed that RL packets is small enough to be transmitted in one multi-frame i.e. $T_{Transmission}$ = 60 milliseconds. Under the assumption that the packet error rate is less than 5% the 95% percentile of TX_{RL} will then be below

---

[10]This approach neglects the fragmentation of large higher layer packets. However, as the most stringent latency requirements apply to the smallest packets, it is suitable to provide a valid estimate of the minimum DC slot size.

[11]Sending the allocation in the next CC is not strictly required in the specification, but strongly recommended in the LDACS1 guidance material. If the system is not overloaded there is, however, no reason to delay the allocation anyway.

$$T_{RL95} \leq \left\lceil \frac{AS}{DC_{AS}} \cdot MF \right\rceil + 2 \cdot MF \tag{3}$$

as 95% of the packets will be successfully transmitted in this time. The 5% erroneous packets require additional time for a retransmission.

Putting the equations together the DC slot length required to meet a 95% percentile requirement L can be calculated from:

$$T_{RL95} \leq \left\lceil \frac{AS}{DC_{AS}} \cdot MF \right\rceil + 2 \cdot MF \leq L \tag{4}$$

The minimum DC slot length to meet the requirement L with AS aircraft stations is thus:

$$DC_{AS} \geq \frac{MF \cdot AS}{L - 2 \cdot MF} \tag{5}$$

This formula can now be used to derive the minimum DC slot length necessary to support a given 95% latency requirement. The results of this calculation for the evaluation scenarios of Section 4.1 and the requirements in Table 4 of section 4.3 are displayed in Table 1. Note that all minimum slot lengths are below the maximum physical layer DC slot size of 52 tiles i.e. the formal analysis indicates that the LDACS1 medium access sub-layer design scales to fulfil the defined latency requirements.

Scenario[12]	Number of aircraft (PIAC)	Minimum DC slot size (tiles)			
		ATS Only, with A-EXEC	ATS+AOC, with A-EXEC	ATS Only, without A-EXEC	ATS+AOC, without A-EXEC
APTZone	26	-	-	2	2
APTSurface	264	-	-	13	13
TMASmall	44	5	5	3	3
TMALarge	53	6	6	3	3
ENRSmall	45	5	5	3	3
ENRMedium	62	6	6	3	3
ENRLarge	204	20	20	10	10
ENRSuperLarge	512	50	50	24	24

Table 1. LDACS1 minimum DC slot size in tiles.

3.5 Logical link control sub-layer

The logical link control sub-layer contains the necessary protocols to facilitate reliable data transfer for multiple users. It comprises the Data Link Service (DLS) and the Link Management Entity (LME).

---

[12]Scenarios are discussed in Section 4.1.

3.5.1 Data link service analysis

The DLS has two major functions: First the segmentation and reassembly of higher layer packets. Second the acknowledged and unacknowledged transmission of (fragmented) higher layer packets. In addition, the DLS performs the local quality of service management using separate queues for different service classes. If a resource grant is received, the input queues are served according to their priorities. Higher priority traffic classes may pre-empt lower priority classes. This guarantees high priority queues to get prioritised medium access.

The design of the reliable DLS protocol had to be carried out taking two main requirements into consideration: Reliability and responsiveness. Reliability is formalized by the notion of continuity i.e. it is linked with the detection and recovery of lost and duplicated packets. The analysis below indicates under which conditions ARQ is required to ensure the stated continuity requirements and justifies the use of ARQ in LDACS1.

High levels of continuity can either be achieved by the application of error correcting codes (Forward Error Correction; FEC), the retransmission of erroneous packets (Automatic Repeat request; ARQ; sometimes also called backward error correction), or a combination of both approaches (Hybrid ARQ; HARQ). The continuity c that can be achieved using these approaches can be calculated by

$$c(p,R) = \sum_{n=0}^{R} (1-p)\, p^n \tag{6}$$

if the higher layer message is small enough to be conveyed in a single packet. The variable p is the effective packet error rate after FEC (if forward error correction is applied). The variable R is the maximum number of retransmissions supported by the ARQ protocol. Note that $R=0$ is equivalent to not using ARQ.

For large higher layer messages that need to be transmitted in m fragments continuity is given by

$$c(p,R,m) = c(p,R)^m \tag{7}$$

This analytical model can now be applied to the data link services of AP17 technical task 2 on which the requirements of section 4.3 are based. Fig. 8 plots the achievable continuity for these services without using ARQ. The dashed horizontal line denotes the continuity requirement of the investigated services (99.96%, 99.996%, and 99.999992% in Fig. 8 (a), (b), and (c), respectively), the dashed vertical line indicates the expected bit error rate of the LDACS1 physical layer (10^{-5}) i.e. the continuity requirement is fulfilled by LDACS1 if the continuity curve of each COCRv2 service intersects with the dashed horizontal line right of the dashed vertical line. The minimum DLS-PDU size of 125 byte is assumed (cf. section 4.2). The expected bit error rate of LDACS1 after FEC is 10^{-5}. This is indicated with vertical dotted lines. The required level of continuity is marked by the horizontal dotted line (i.e. 99.9x%). Note that different service classes have differing requirements.

The calculations show clearly that the required continuity cannot be achieved with the current FEC and without ARQ. Not allowing retransmissions the effective frame error rate would have to be decreased by three orders of magnitude.

The usage of strong error correcting codes in LDACS1 cannot be avoided in any case due to the unfavourable interference conditions of the radio channel. Using the default coding and modulation a bit error rate of 10^{-5} can be achieved using FEC. However, in this case the coding rate is approximately ½ i.e. for each bit of information two bits have to be

transmitted on the channel. Lower frame error rates would come at the cost of over-proportionally increasing the coding overhead even more, thus any remaining errors have to be recovered by retransmissions i.e. LDACS1 has to apply HARQ[13].

(a)

(b)

(c)
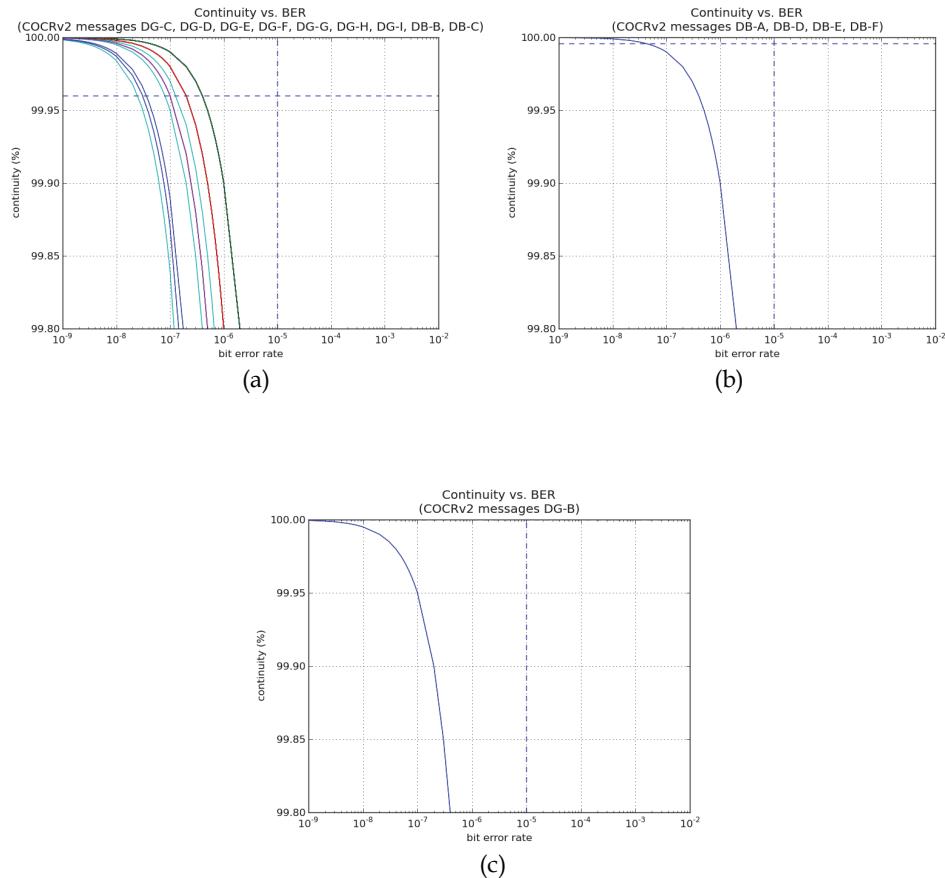
Fig. 8. Continuity of COCRv2 messages (no ARQ).

---

[13]The precise term is Type 1 HARQ. The more advanced Type 2 HARQ (with progressively increasing FEC) is not used in LDACS1.

(a)

(b)

(c)
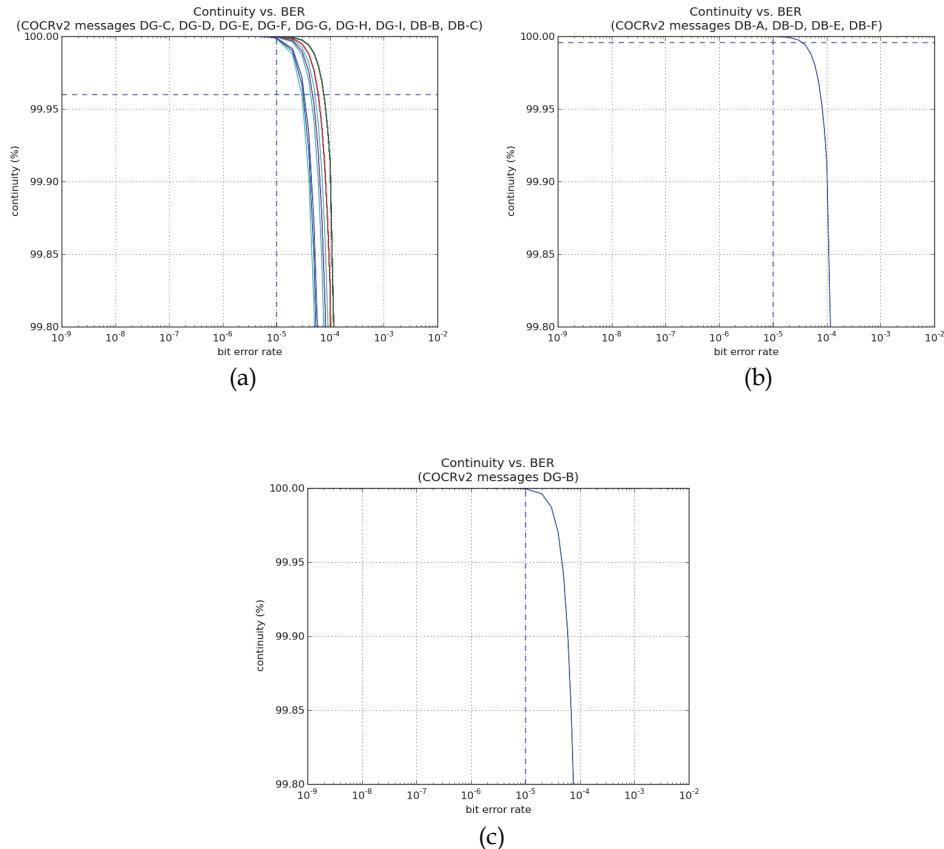
Fig. 9. Continuity of COCRv2 messages (ARQ up to 2 retransmissions).

Fig. 9 displays the levels of continuity that can be achieve by allowing up to two ARQ retransmissions (R=2). The minimum DLS-PDU size of 125 byte is assumed i.e. as the packet error rate increases with larger DLS-PDU sizes the identified value for R is a lower bound for the required number of retransmissions. The results indicate that the requirements can now be fulfilled for the expected frame error rate of LDACS1 and for all service classes in this case.

An additional benefit of ARQ is that non-duplication of messages is enforced by the protocol. Thus the main focus of the DLS ARQ design was on the efficient recovery from packet losses through retransmissions. The efficiency of the retransmission mechanism is determined by the ability to retransmit lost fragments of a message such that the quality of service requirements of the message are not violated i.e. the protocol had to be designed to provide quick retransmissions in order to be effective.

3.5.2 Data link service timer management

DLS timer management was identified as crucial for the overall protocol performance in the LDACS1 evaluation. It was therefore proposed to couple the DLS timer management to the

MAC time framing to achieve near optimal performance. Thus, the LDACS1 ARQ protocol operates in fixed timing relations to the medium access sub-layer.

Depending on the physical layer implementation (i.e. decoding latency) there are two optimal acknowledgement opportunities for a DLS-PDU transmitted in a Data slot: The first opportunity is to acknowledge the DLS-PDU in the control channel slot of the same multi-frame. The second opportunity is to use the first control channel slot after the current multi-frame (e.g. if the transmission of the DLS-PDU and the control channel slot overlap).

The retransmission timer defines the maximum number of missed acknowledgement opportunities, before a retransmission is triggered. According to the last paragraph, the DLS retransmission timer should time out after not more than two missed acknowledgement opportunities.

Fig. 10 illustrates the DLS retransmission timer on the reverse link. After the aircraft-station has sent a DLS-PDU in the RL Data slot there are two possible acknowledgment opportunities on the forward link. The DLS retransmission timer is set to the end of the second acknowledgement opportunity.
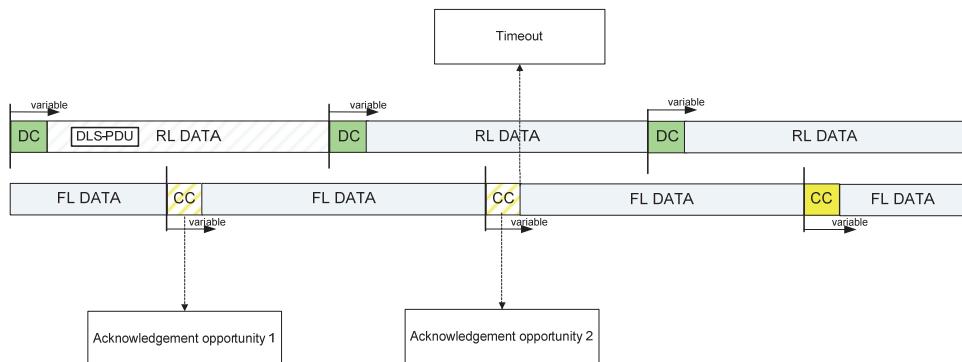

Fig. 10. LDACS1 RL DLS retransmission timer.
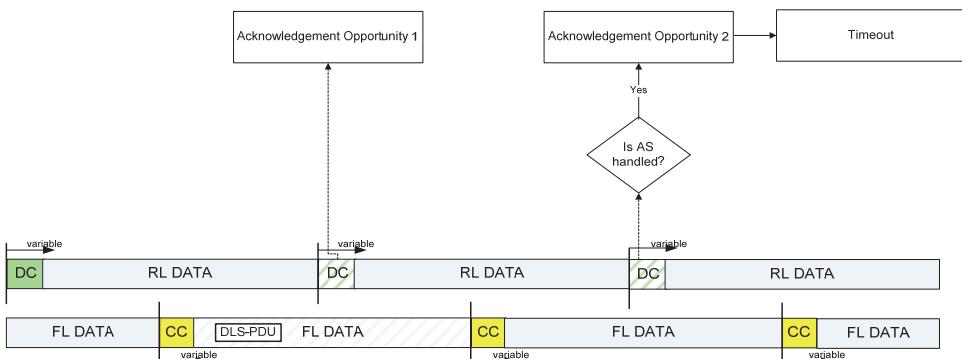
Fig. 11. LDACS1 FL DLS retransmission timer.

Fig. 11 displays the same concept for the forward link DLS retransmission timer. There are two acknowledgement opportunities on the reverse link after the DLS Data PDU is sent. The first opportunity could be in the DC slot during the FL Data slot of the transmission. The second acknowledgement opportunity is the next appearance of the receiving aircraft-station's DCCH. According to the number of registered aircraft-stations not every aircraft-station is able to send its DCCH in each multi-frame. The timeout is therefore set to the end of the next DC slot where the aircraft-stations DCCH is transmitted.

Note that the first DC slot can always be counted as an acknowledgement opportunity, as the acknowledgement has to be sent in the next DC slot in any case.

3.5.3 Link management entity resource allocation

Channel resources for transmission have to be requested by the DLS from the radio resource management in the ground-station LME. The ground-station DLS makes these requests locally using an internal interface, while the aircraft-station DLS has to make its request over the dedicated control channel DCCH. The radio resource management stores these requests to calculate an appropriate resource allocation. Note that the aircraft-station DLS makes separate resource requests for each class of service. Requests are encoded in a single (variable length) control message.

The resource allocation for the next forward link and reverse link data slots is then calculated after the end of the DC slot when the LME has collected the resource requests of all aircraft-stations serviced in this multi-frame. The gap between the DC slot and the CC slot is used to calculate the assignment and to prepare it for transmission in the CC (i.e. generate FL and RL allocation control messages; cf. Fig. 6).

The allocation algorithm has not been defined in the LDACS1 specification, but is left open to the implementer. The simulations presented in this chapter use a comparatively simple prioritized round-robin resource allocation algorithm. This algorithm respects the priorities of the different classes of service and is fair between requests of the same priority.

The resource allocation of the forward link and reverse link are calculated independently, but following the same approach: In the first step the resource requests are sorted according to their priority and class of service: High priorities before low priorities, and acknowledged transmissions before unacknowledged transmissions. In the second step resource allocations are granted in round-robin within each class of service. If the class has been completely serviced the next class is served. The algorithm assigns the largest possible allocation limited by the size of the data slot.

For the reverse link allocations an additional restriction is introduced to reduce the duty cycle and the interference generated by the system. The maximum size of the resource allocation is limited to 16 OFDM tiles. This is equivalent to a maximum reverse link sending time of 5.76 milliseconds. Note that this restriction has no formal motivation. It was assumed as a working hypothesis in the physical layer development. Its size can, however, be configured.

Note that only the sum of all resource allocations is transmitted to the user as it can be locally distributed by the DLS quality of service function according to the transmitted requests.

4. Design validation

The design validation of the LDACS1 protocol was carried out in a computer simulation implementing the medium access sub-layer and the logical link control sub-layer. This section presents the discussion of selected simulation results according to the evaluation criteria stated in Section 2.1 and formalized in section 4.3.

4.1 Simulation scenarios

The design validation was performed on the basis of the data traffic profile (also called "mobile communication operational concept") defined in the COCRv2 report (EUROCONTROL & FAA, 2007b) and the air traffic volumes defined in the companion document to the COCRv2 (EUROCONTROL & FAA, 2007c). Both documents were produced in AP17 technical task 2.

The COCRv2 mobile communications operational concept is very detailed and is described on the basis of a set of anticipated (i.e. hypothetical) data link services. These services are divided into three categories: ATC, AOC, and Network Management Services (NET). Each of these services is described in great detail: The events triggering the generation of a data packet, size and quantity of the packet, expected reaction of the peer entity (i.e. responses or acknowledgements), and the expected class of service.

It was recognized by the authors of the COCRv2 report that this elaborate model is non-trivial to implement, in particular, as it requires a detailed air traffic model. Therefore the COCRv2 report was augmented with the companion document (EUROCONTROL & FAA, 2007c) containing a set of simplified evaluation scenarios. These simplified scenarios were designed such as to ensure that all COCRv2 services can be supported if the simplified requirements are fulfilled i.e. they are worst case scenarios.

These evaluation scenarios are "simplified" in the sense that the scenarios of the companion document are easier implemented by the evaluator. The simplified scenarios were also created using the COCRv2 data link service descriptions. However, they were already based on synthetic air traffic situations (i.e. artificial air traffic provided by the authors of the COCRv2 report) referred to as "air traffic volumes".

An excerpt of the relevant air traffic volumes is cited in Table 2. The "APT Surface" traffic volume has been left aside in this section as this communication domain is covered by the dedicated IEEE 802.16e based airport surface data link. The "ENR Super Large" traffic volume covers an area larger than the area that could be covered using the theoretical maximum range of L-DACS1, which is 200 nautical miles. It is therefore left aside, too.

Except for the APT Zone traffic volume, which is cylindrical, all traffic volumes are cuboids of different sizes. The TMA and ENR traffic volumes have constant heights of 19,500 feet (approximately 6,000 meters) and 20,500 feet (approximately 6,300 meters), respectively. Each traffic volume has a Peak Instantaneous Aircraft Count value (PIAC) reporting the maximum number of aircraft to be expected within this volume at the same time.

In addition to the introduction of synthetic air traffic, data traffic is no longer presented at the packet level, but aggregated into average user data rates (i.e. offered load) as displayed in Table 3. The user data rates are split into four scenarios: Either ATS traffic alone or ATS and AOC traffic combined, with or without the very demanding A-EXEC[14] service. The most challenging scenario is the ATS+AOC scenario with A-EXEC service.

---

[14]"The A-EXEC service provides an automated safety net to capture situations where encounter-specific separation is being used and a non-conformance FLIPINT event occurs with minimal time remaining to

Ref.	Type	Dimensions	Height Range	Number of aircraft (PIAC)
TV 1.1	APT Zone	Cylinder, 10 NM[15] diameter	0 – FL50[16]	26
TV 1.2	APT Surface	Cylinder, 5 NM diameter	0	264[17]
TV 2.1	TMA Small	Cuboid, 49 x 49 NM	FL50 – FL245	44
TV 2.2	TMA Large	Cuboid, 75.0 x 75.0 NM	FL50 – FL245	53
TV 3.1	ENR Small	Cuboid, 55 x 55 NM	FL245 – FL450	45
TV 3.2	ENR Medium	Cuboid, 100.0 x 100.0 NM	FL245 – FL450	62
TV 3.3	ENR Large	Cuboid, 200.0 x 200.0 NM	FL245 – FL450	204
TV 3.4	ENR Super Large	Cuboid, 400.0 x 400.0 NM	FL245 – FL450	522

Table 2. Traffic volumes. Cited from (EUROCONTROL & FAA, 2007c).

At the start of each simulation scenario the air traffic volume is populated with aircraft. The scenarios do not require the simulation of cell entry and cell exit therefore it is assumed that the number of aircraft is constant at the PIAC during the complete simulation time. The position of the aircraft within the traffic volume is not simulated as the physical layer simulations covered the influence of the aircraft's position already with worst case assumptions.

4.2 Simulation parameters

The simulation implemented the LDACS1 protocol according to (Sajatovic et al., 2011), and the descriptions given in the previous sections.

The evaluation scenarios were simulated using a higher layer packet size of 125 Byte[18] with a single class of service[19]. Each higher layer packet is extended by a 3 Byte SNDCP header and at least one 7 Byte DLS-PDU header (assuming minimal fragmentation). Together these headers add 10 Bytes of overhead to each higher layer packet which is equivalent to 8%

---

resolve the conflict. […] When non-conformance occurs, triggering an imminent loss of separation, the ground automation system generates and sends a resolution to the aircraft for automatic execution without the Flight Crew or Controller in the loop." (EUROCONTROL & FAA, 2007b)

[15]Nautical miles.

[16]Flight Level FL expresses the aircraft altitude (above mean sea level) in steps of 100 ft (e.g. FL50 corresponds to an aircraft altitude of 5000 ft above mean sea level). Note that the abbreviation FL is also used for the Forward Link FL (i.e. ground-to-air) transmission direction depending on the context.

[17]The APT Surface traffic volume contains all aircraft on the ground. The other traffic volumes contain only airborne aircraft.

[18]Note that this is the near the most common packet size defined in COCRv2 and that the SNDCP may perform additional fragmentation and reassembly if necessary.

[19]Simulation results for larger packet sizes and several service classes can be found in (Gräupl et al., 2009).

overhead. The simulation duration was set to 500 seconds plus 5 additional seconds of follow up time. Each scenario was simulated ten times with different random seeds.

Scenario	PIAC	Average User Data Rate (kbit/s)							
		ATS Only, with A-EXEC		ATS + AOC, with A-EXEC		ATS Only, without A-EXEC		ATS + AOC, without A-EXEC	
		FL[20]	RL	FL	RL	FL	RL	FL	RL
APT Zone	26	-	-	-	-	3	10	3	15
APT Surface	264	-	-	-	-	30	30	150	30
TMA Small	44	30	30	30	30	30	30	30	30
TMA Large	53	30	30	30	30	30	30	30	30
ENR Small	45	30	30	150	30	30	30	80	30
ENR Medium	62	30	30	150	30	30	30	100	30
ENR Large	204	30	40	300	40	30	40	200	40
ENR Super Large	522	40	50	500	50	40	50	500	50

Table 3. Validation scenarios.

All simulation scenarios used the same medium access sub-layer and logical link control sub-layer settings. If not stated otherwise the DC slot size was set to 52 tiles. The maximum reverse link allocation size was set to 16 PHY-SDUs per aircraft-station and multi-frame. This is equivalent to a maximum reverse link sending duration of 5.76 milliseconds per multi-frame not taking the DC slot into account. The DLS ARQ window size was set to 4 DLS-SDUs with up to 4 DLS-PDUs per transmission buffer. The maximum DLS-PDU size was set equal to the DLS-SDU size (i.e. no fragmentation).

The physical layer bit error rate was set to 10-5 after FEC. Bit errors were simulated for all logical channels. With the exception of the "ENR Large ATS+AOC with A-EXEC" scenario the default coding and modulation was used. This scenario uses ACM type 2 (QPSK, coding rate 2/3) coding. The "ENR Super Large ATS+AOC with A-EXEC" and "ENR Super Large ATS+AOC with-out A-EXEC" are actually out of the scope of LDACS1, but were included into the simulations with ACM type 4 coding and modulation (16QAM, coding rate 1/2). Results using non-default ACM settings are indicated in *italic* font.

4.3 Validation parameters

The LDACS1 design goals were validated according to the evaluation scenarios defined above and the quality of service requirements defined in Section 4.6 of (EUROCONTROL & FAA, 2007c). The quality of service requirements are defined in terms of continuity and the 95% percentile of the one-way transmission latency (TT$_{95}$-1 way).

All values of interest are estimated by the mean of ten measurements. The measurements considered in the presented results are defined as follows:

Latency: The one-way latency of a data packet is calculated as the time between the creation of the packet and its successful reception. This value is validated against the "TT95-1 way" requirement.

---

[20]In this table FL and RL denote Forward Link (ground-to-air) and Reverse Link (air-to-ground) directions.

Continuity: The percentage of packets that is not lost, duplicated or expired. This value is validated against the "Continuity" requirement in Table 4.

Scenario	Without A-EXEC		with A-EXEC	
	TT95-1 way (s)	Continuity (%)	TT95-1 way (s)	Continuity (%)
APT	1.4	99.96	-	99.999992
TMA	1.4	99.96	0.74	99.999992
ENR	1.4	99.96	0.74	99.999992
ORP	5.9	99.96	-	99.999992
AOA	1.4	99.96	-	99.999992

Table 4. Quality of service requirements. Cited from (EUROCONTROL & FAA, 2007c).

4.4 Results

The design goal for responsiveness is fulfilled if the 95%-percentile values of the LDACS1 one-way latency satisfy the requirements of Table 4. The results presented in Table 5 indicate that LDACS1 fulfils all these requirements in the presented validation scenarios. Note that the "ENR Large ATS+AOC with A-EXEC" scenario uses ACM type 2. "ENR Super Large ATS+AOC with A-EXEC" and "ENR Super Large ATS+AOC with-out A-EXEC" use ACM type 4. The PIAC in these scenarios was changed from 522 aircraft to 512 aircraft as this is the maximum supported by LDACS1 in a single radio cell.

The latency requirements of the A-EXEC service (740 milliseconds) is fulfilled in all cases. This indicates that the DC slot size (and hence the dedicated control channel capacity) may be reduced. Table 6 displays the 95% percentile results for reduced dedicated control channel capacity. The size of the DC slot was constrained to the minimum number of tiles according to Table 1.

The results show that the 95% percentile of the one-way latency changes indeed as predicted. With the exception of the ENR Medium no-A-EXEC scenario (DC is 3 tiles, TT_{95}-1 way requirement is 1400 milliseconds) all requirements are met with the minimum slot sizes. The failed requirement poses no real problem as the DC slot size can be easily increased above the minimum value. It should be noted that the forward link latency also increases when the DC slot is smaller. This is caused by the ARQ transmission window which can only be shifted at the reception of a new acknowledgement in the dedicated control channel.

The results indicate that the theoretical analysis of section 3.4.2 provides a good starting point for optimization but tends to underestimate the required DC slot size in realistic scenarios. However, the requirements are only missed by a small margin (72 ms).

Scenario	PIAC	95% percentile of latency (TT95-1 way)							
		ATS Only, with A-EXEC		ATS+AOC, with A-EXEC		ATS Only, without A-EXEC		ATS+AOC, without A-EXEC	
		FL	RL	FL	RL	FL	RL	FL	RL
APT Zone	26	-	-	-	-	125	412	126	412
APT Surface	264	-	-	-	-	134	178	134	179
TMA Small	44	128	180	128	180	128	180	128	180
TMA Large	53	125	187	125	187	125	187	125	187
ENR Small	45	127	180	128	180	127	180	127	180
ENR Medium	62	125	227	126	227	125	227	125	227
ENR Large	204	125	350	*161*	*349*	125	350	129	350
ENR Super Large	512	125	695	*212*	*693*	125	695	*212*	*693*

Table 5. LDACS1 responsiveness (TT95-1 way); DC size 52.

Scenario	PIAC	95% percentile of latency (TT95-1 way)							
		ATS Only, with A-EXEC		ATS+AOC, with A-EXEC		ATS Only, without A-EXEC		ATS+AOC, without A-EXEC	
		FL	RL	FL	RL	FL	RL	FL	RL
APT Zone	26	-	-	-	-	141	868	145	868
APT Surface	264	-	-	-	-	139	1296	141	1296
TMA Small	44	143	639	143	639	143	988	143	988
TMA Large	53					144	1146	144	1146
ENR Small	45	142	646	430	635	144	994	579	996
ENR Medium	62	143	724	334	719	144	1321	1472	1323
ENR Large	204	137	706	*187*	*708*	141	1298	218	1307
ENR Super Large	512	126	709	*207*	*711*	136	1350	*272*	*1350*

Table 6. LDACS1 responsiveness (TT95-1 way) ; minimum DC size.

Scenario	PIAC	Continuity in %							
		ATS Only, with A-EXEC		ATS+AOC, with A-EXEC		ATS Only, without A-EXEC		ATS+AOC, without A-EXEC	
		FL	RL	FL	RL	FL	RL	FL	RL
APT Zone	26	-	-	-	-	100	100	100	100
APT Surface	264	-	-	-	-	100	100	100	100
TMA Small	44	100	100	100	100	100	100	100	100
TMA Large	53	100	100	100	100	100	100	100	100
ENR Small	45	100	100	100	100	100	100	100	100
ENR Medium	62	100	100	100	100	100	100	100	100
ENR Large	204	100	100	*100*	*100*	100	100	100	100
ENR Super Large	512	100	100	*100*	*100*	100	100	*100*	*100*

Table 7. LDACS1 continuity ; DC size 52.

4.4.2 Reliability

The evaluation of the LDACS1 continuity in the defined simulation scenarios shows that LDACS1 can fulfil the continuity requirements of (EUROCONTROL & FAA, 2007c) in all cases.

4.4.3 Scalability

The fact that LDACS1 fulfils the COCRv2 requirements in all investigated cases indicates that the system provides the required scalability.

5. Conclusion

The objective of the LDACS1 development was to create a first protocol specification enabling prototyping activities. It was not the goal of this development to create a final product and it is expected that further refinements of the protocol will originate from prototyping. However, the analysis, design, and validation of LDACS1 produced a framework of protocols backed by formal and simulation based analysis. The goal was to develop a protocol design providing the quality of service required for future ATM operations.

The LDACS1 research produced a deterministic medium access approach built on the lessons learnt from its predecessor protocols. This approach ensures that the medium access latency is only coupled to the number of aircraft-stations served by the ground-station. The medium access performance degrades only linearly with the number of users and not exponentially as in the case of random access. In the LDACS1 protocol design the resource allocation between different users is performed centralized by the ground-station while the

resource distribution between packets of different priorities is performed locally by each user. The effect of this approach is that the medium access sub-layer supports prioritized channel access.

The analysis of the requirements towards the overall communication system performance produced the justification for the use of ARQ in the LDACS1 logical link control sub-layer. Coupling the DLS timer management to the MAC sub-layer time framing has the effect to produce near to optimal timer management. LDACS1 can thus be considered a mature technology proposal offering a solid baseline for the definition of the future terrestrial radio system envisaged in AP17.

LDACS1 has now entered a new phase within the protocol engineering process going from the development phase to the prototyping phase. The initial specification can now be considered complete and evaluated. The next steps will be determined by the further optimization of the protocol and the evaluation of the prototype within the context of the Single European Sky ATM Research Programme (SESAR).

6. References

Brandes, S.; Epple, U.; Gligorevic, S.; Schnell, M.; Haindl, B. & Sajatovic, M. (2009). Physical Layer Specification of the L-band Digital Aeronautical Communications System (L-DACS1), Proceedings ICNS'09, ISBN 978-1-4244-4733-6, Washington DC, May 2009.

Budinger, J. & Hall, E. (2011). Aeronautical Mobile Airport Communications System (AeroMACS), In: *Future Aeronautical Communications*, Plass, S., InTech, ISBN 979-953-307-443-5

Commision of the European Communities. (2001). European transport policy for 2010: time to decide, Office for official publications of the European Communities, ISBN 92-894-0341-1, Brussels

Eleventh Air Navigation Conference. (2003). Report of Committee B to the Conference on Agenda Item 7, Availabe from: http://www.icao.int/icao/en/anb/meetings/anconf11/documentation/anconf11_wp202_en.pdf

EUROCONTROL & FAA. (2007a). Action Plan 17 Future Communications Study - Final Conclusions and Recommendations, Available from: http://www.eurocontrol.int/communications/gallery/content/public/documents/AP17_Final_Report_v11.pdf

EUROCONTROL & FAA. (2007b). Communication Operating Concept and Requirements for the Future Radio System, Ver. 2, Available from: http://www.eurocontrol.int/communications/gallery/content/public/documents/COCR%20V2.0.pdf

EUROCONTROL & FAA. (2007c). Evaluation Scenarios, Available from: http://www.eurocontrol.int/communications/gallery/content/public/documents/FCS_Eval_Scenarios_V10.pdf

European Commission. (2011). Single European Sky, Available from: http://ec.europa.eu/transport/air/single_european_sky/single_european_sky_en.htm

Fistas, N. (2009). Future Aeronautical Communication System – FCI, Proceedings of Take Off Conferenece, Salzburg, April 2009.

Gräupl, T.; Ehammer, M.; & Rokitansky, C.-H. (2009). LDACS1 Data Link Layer Design and Performance, Proceedings of ICNS'09, ISBN 978-1-4244-4733-6, Washington DC, May 2009.

Haindl, B.; Rihacek, C.; Sajatovic, M.; Phillips, B.; Budinger, J.; Schnell, M.; Lamiano, D. & Wilson, W. (2009). Improvement of L-DACS1 Design by Combining B-AMC with P34 and WiMAX Technologies, Proceedings of ICNS'09, ISBN 978-1-4244-4733-6, Washington DC, May 2009.

Helfrick, A. (2007). Principles of Avionics (4th ed.). Airline Avionics, ISBN 978-1885544261, Leesburg, VA

IATA. (2003). IATA Position on Aeronautical Air Ground Communications Needs, Available from:
http://www.icao.int/icao/en/anb/meetings/anconf11/documentation/ANConf11_wp054_en.pdf

Kamali, B. (2010). An Overview of VHF Civil Radio Network and the Resolution of Spectrum Depletion, Proceedings of ICNS'10, ISBN 2155-4943, Washington DC, May 2010.

Sajatovic, M.; Haindl, B.; Epple, U. & Gräupl, T. (2011). Updated LDACS1 System Specification. SESAR P15.2.4 EWA04-1 task T2 Deliverable D1.

Rokitansky, C.-H.; Ehammer, M.; Gräupl, T.; Schnell, M.; Brandes, S.; Gligorevic, S.; Rihacek, C. & Sajatovic, M. (2007). B-AMC a system for future broadband aeronautical multi- carrier communications in the L-band, Proceedings of 26th DASC, ISBN 978-1-4244-1108-5, Dallas TX, Nov. 2007.

The LDACS1 Physical Layer Design

Snjezana Gligorevic, Ulrich Epple and Michael Schnell
German Aerospace Center (DLR)
Oberpfaffenhofen,
Germany

1. Introduction

The legacy DSB-AM (Double Sideband Amplitude Modulation) system used for today's voice communication in the VHF-band is far away of meeting the demands of increasing air traffic and associated communication load. The introduction of VDL (VHF Digital Link) Mode 2 in Europe has already unfolded the paradigm shift from voice to data communication. Legacy systems, such as DSB-AM and VDL Mode 2 are expected to continue to be used in the future. However, they have to be supplemented in the near future by a new data link technology mainly for two reasons. First, only additional communication capacity can solve the frequency congestion and accommodate the traffic growth expected within the next 10-20 years in all parts of European airspace (ICAO-WGC, 2006). Second, the modernization of the Air Traffic Management (ATM) system as performed according to the SESAR (http://www.sesarju.eu/) and NextGen (http://www.faa.gov/nextgen/) programs in Europe and the US, respectively, heavily relies on powerful data link communications which VDL Mode 2 is unable to support.

Based on the conclusions of the future communications study (Budinger, 2011), the ICAO Working Group of the Whole (ICAO-WGW, 2008) has foreseen a new technology operating in the L-band as the main terrestrial component of the Future Communication Infrastructure (FCI) (Fistas, 2011) for all phases of flight. Hence, such L-band technology shall meet the future ATM needs in the en-route and the Terminal Manoeuvring Area (TMA) flight domains as well as within airports. The latter application area will be supplemented by the AeroMACS technology at many large airports (Budinger, 2011).

A final choice of technology for the L-band has not been made yet. Within the future communications study, various candidate technologies were considered and evaluated. However, it was found that none of the considered technologies could be fully recommended before the spectrum compatibility between the proposed systems and the legacy systems has been proven. This will require the development of prototypes for testing in a real environment against operational legacy equipment.

The future communications study has identified two technology options for the L-band Digital Aeronautical Communication System (LDACS) as the most promising candidates for meeting the requirements on a future aeronautical data link. The first option, named LDACS1, is a Frequency-Division Duplex (FDD) configuration utilizing Orthogonal Frequency-Division Multiplexing (OFDM), a highly efficient multi-carrier modulation technique which enables the use of higher-order modulation schemes and Adaptive Coding and Modulation (ACM). OFDM has been adopted for current and future mobile radio communications technologies,

like 3GPP LTE (Third Generation Partnership Project Long Term Evolution) and 4G (Fourth Generation mobile radio system). In addition, LDACS1 utilizes reservation based access control (Gräupel & Ehammer, 2011) to guarantee timely channel access for the aircraft and advanced network protocols similar to WiMAX (Worldwide Interoperability for Microwave Access) and 3GPP LTE to ensure high quality-of-service management and efficient use of communication resources. LDACS1 is closely related to the Broadband Aeronautical Multi-Carrier Communication (B-AMC) and TIA-902 (P34) technologies (Haindl at al., 2009).

LDACS2 is the second option which is based on a single-carrier technology. It utilizes a binary modulation derivative (Continuous-Phase Frequency-Shift Keying, CPFSK) and thus does not enable the use of higher-order modulation schemes. For duplexing Time-Division Duplex (TDD) is chosen. The physical layer has some similarities to both the Universal Access Transceiver (UAT) and the second generation mobile radio system GSM (Global System for Mobile Communications). A custom protocol is used providing high quality-of-service management capability. This option is a derivative of the L-band Data Link (LDL) and the All-purpose Multi-channel Aviation Communication System (AMACS) technologies (EUROCONTROL, 2007).

Follow-on activities required in order to validate the performance of the proposed LDACS options and their compatibility with legacy L-band systems, finally aiming at a decision on a single L-band technology, run under the SESAR framework (http://www.sesarju.eu/; Fistas, 2011).

2. System requirements

The choice of the radio link is based on the capacity the link should provide related primarily to the services and applications that it should support. The radio frequency will affect the propagation loss, whereas the channel fading in a deterministic environment may also vary with the system bandwidth. Additionally, the interference conditions in the part of the L-band assigned to the Aeronautical Mobile (Route) Service (AM(R)S) have to be considered. Consequently, the development of an air-ground data link in the L-band faces several requirements, both operational and technical.

2.1 Services and applications

Air Traffic Services (ATS) and Airline Operational Communications (AOC) services are related to safety and regularity of flight and hence entail more stringent requirements on a future communication system in comparison with commercial mobile communication systems.

One of the requirements for a new data link in the L-band is the suitability to support future services and applications as described in (EUROCONTROL & FAA, 2007). The document describes safety, information security, and performance assessments for the air traffic services, derives high-level requirements that each service would have to meet and allocates the requirements to the future radio system. Beside a range of parameters on which the suitability of communication systems can be assessed, the document provides capacity requirements estimated for different service volumes and regarding increasing air traffic and future communication concepts.

2.2 Propagation conditions

Typically, during the flight an aircraft traverses numerous Air Traffic Control (ATC) sectors and en-route facilities. In comparison to the VHF band used by the legacy ATC systems,

higher free space loss in the L-band implies smaller sector sizes. The possibility of increasing transmitter (Tx) power is limited by the interference constraints and the amplifier dimensions. Hence, the reuse factor of the cellular LDACS system and the interference constraints within the L-band should be taken into account not only for the link budget calculation but also for frequency planning for the European airspace.

Furthermore, the sector size affects the system capacity in terms of data throughput per aircraft, but also the system design in terms of required guard times between Forward Link (FL) and Reverse Link (RL). Whereas in the FDD configuration, as for LDACS1, the guard times have to be guaranteed only in the random access phase, the general requirement for guard times in a TDD based system implies a loss in the system capacity.

In en-route domain, propagation conditions are characterized by a very strong Line-Of-Sight (LOS) component, and thus, multipath effects have only very limited influence on the received signal quality. More severe multipath conditions in the TMA and airport domains result in increased frequency selectivity of the channel. A broadband system may benefit from the frequency diversity related to the multipath, whereas a narrowband system will be affected by more severe fading on the LOS path between transmitter and receiver.

According to the publications on propagation conditions in L-band based on measurements (Rice et al, 2004) and on theoretical considerations (ICAO-WGC, 2006), the Root Mean Square Delay Spread (RMS-DS) remains below 2 µs in en-route case. The maximum delay and delay spread increase in TMA and airport areas. Measurements at airports provide a maximum RMS-DS of 4.5 µs and 90th percentile delay spread not exceeding 1.7 µs during taxiing (Gligorevic et al., 2009; Matolak et al., 2008).

Taking into account an aircraft velocity of 1050 km/h in an en-route area we obtain a maximum Doppler frequency of 972 Hz. However, due to the dominant LOS component in en-route domain, the Doppler effect will mainly cause a Doppler shift of the carrier frequency. Since the velocity is lower in TMA and especially in airport areas, the Doppler spread resulting from the Doppler effect in the reflections of the signal will be lower. According to (Bello, 1973), the reflections in the L-band can be modelled as a Rayleigh process with a Gaussian Doppler spectrum.

2.3 Spectrum

LDACS shall operate in the lower part of the L-band, 960 - 1164 MHz. As depicted in Fig. 1, the L-band is already utilized by several systems.

The Distance Measuring Equipment (DME) operating as an FDD system on the 1 MHz channel grid is a major user[1] of the L-band. Parts of this band are used in some countries by the military Multifunctional Information Distribution System (MIDS). Several fixed channels are allocated for the Universal Access Transceiver (UAT) and the Secondary Surveillance Radar (SSR)/Airborne Collision Avoidance System (ACAS). Fixed allocations have been made in the upper part of the L-band for the Global Position System (GPS), Global Orbiting Navigation Satellite System (GLONASS) and GALILEO channels. The commercial mobile radio systems UMTS (Universal Mobile Telecommunications System) and GSM are operating immediately below the lower boundary of the aeronautical L-band (960 MHz). Additionally, different types of RSBN (Радиотехническая система ближней навигации is a Russian air navigation system)

---

[1]DME channels are also used by the military Tactical Air Navigation (TACAN) system.

may be found in some parts of the world, operating on channels 960 - 1164 MHz (SESAR JU, 2011).
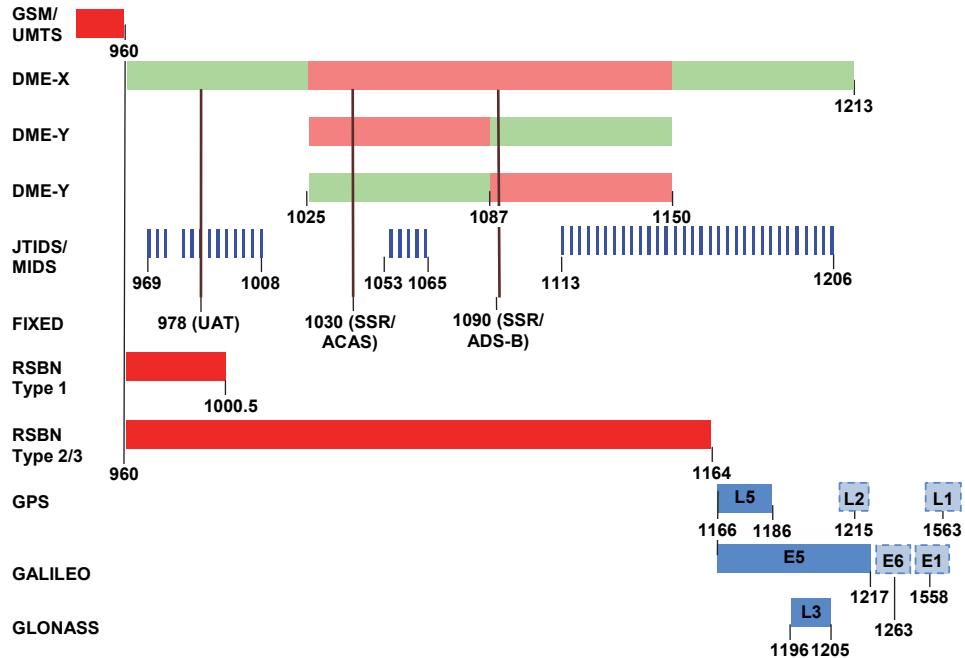
Fig. 1. Current L-band usage (SESAR JU, 2011).

The DME-free part of the spectrum is only between 960 – 975 MHz. Both LDACS systems can use this spectrum of 15 MHz proving not to interfere with the adjacent GSM and UMTS in the lower band, UAT at 978 MHz, and ground DME above 978 MHz. Whereas LDACS2 is expected to operate in the 960-975 MHz frequency band, LDACS1 offers also the opportunity to use spectral gaps between existing DME channels, thus increasing the potential number of communication channels. In this inlay deployment option LDACS1 operates at only 500 kHz offset to assigned DME channels as exemplarily shown in Fig. 2.

One of the challenges to build up a cellular system is to find a sufficient number of channels. In case of LDACS1, RL (air to ground) and FL (ground to air) are separated by FDD. When selecting channels for LDACS1, co-location constraints have to be considered for the aircraft equipment. Additionally, the fixed L-band channels at 978, 1030, and 1090 MHz must be sufficiently isolated from LDACS1 channels by appropriate guard bands. To relax co-site interference problems for an airborne LDACS1 receiver (Rx) in the inlay deployment option, the frequency range 1048.5 - 1171.5 MHz, which is currently used by airborne DME interrogators, should be used for the RL, i.e. airborne LDACS1 Tx. The proposed sub-range for the FL is 985.5 - 1008.5 MHz, i.e. at 63 MHz offset to the RL which corresponds to the DME duplex spacing.

The inlay concept offers the clear advantage that it does not require new channel assignments and the existing assignments can remain unchanged. The physical layer design

of LDACS1, described in the following section, accounts primarily for the inlay concept aiming for coexistence with DME system operating on adjacent channels. However, the LDACS1 design also allows a non-inlay or a mixed inlay/non-inlay deployment without any modifications.
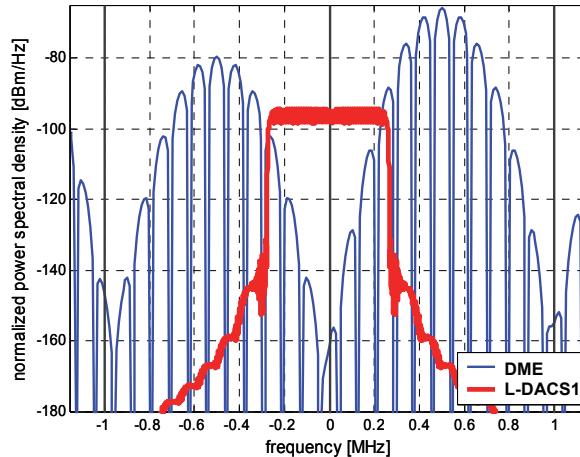

Fig. 2. An example of LDACS1 spectrum and DME interference in the inlay deployment scenario.

3. LDACS1 physical layer characteristics

The LDACS1 physical layer is based on OFDM modulation and designed for operation in the aeronautical L-band (960 – 1164 MHz). Aiming for the challenging inlay deployment option, with limited bandwidth of around 500 kHz available between successive DME channels, and in order to maximize the capacity per channel and optimally use available spectrum, LDACS1 is configured as a FDD system. A TDD approach would be less efficient, since it would require large guard times due to the propagation delays and a split of the available bandwidth into FL and RL transmission. Furthermore, by properly choosing FL and RL frequencies from appropriate parts of the L-band, the co-location interference situation on the aircraft can be significantly relieved.

LDACS1 FL is a continuous OFDM transmission. Broadcast and addressed user data are transmitted on a (logical) data channel, dedicated control and signaling information is transmitted on (logical) control channels. The capacity of the data and the control channel changes according to system loading and service requirements. Message based adaptive data transmission with adjustable modulation and coding parameters is supported for the data channels in FL and RL.

LDACS1 RL transmission is based on Orthogonal Frequency-Division Multiple Access (OFDMA) – Time-Division Multiple Access (TDMA) bursts assigned to different users on demand. In particular, the RL data and the control segments are divided into tiles, hence allowing the Medium-Access Control (MAC) sub-layer of the data link layer to optimize the resource assignments as well as to control the bandwidth and the duty cycle according to the interference conditions.

The channel bandwidth of 498.05 kHz is used by an OFDM system with 50 subcarriers. The resulting subcarrier spacing of 9.765625 kHz is sufficient to compensate a Doppler spread of up to about 1.25 kHz which is larger than typically occurring at aeronautical velocities. For OFDM modulation, a 64-point FFT is used. The total FFT bandwidth comprising all subcarriers is 625.0 kHz.

According to the subcarrier spacing, one OFDM symbol has duration of 102.4 µs. Each OFDM symbol is extended by a cyclic prefix of 17.6 µs, comprising a guard interval of 4.8 µs for compensating multipath effects and 12.8 µs for Tx windowing applied for reduction of the out-of-band radiation. This results in a total OFDM symbol duration of 120 µs. The main LDACS1 OFDM parameters are listed in Table 1.

Parameter	Value
Effective bandwidth (FL or RL)	498.05 kHz
Subcarrier spacing	9.765625 kHz
Used subcarriers	50
FFT length	64
OFDM symbol duration	102.4 µs
Cyclic prefix Guard time Windowing time	17.6 µs 4.8 µs 12.8 µs
Total OFDM symbol duration	120 µs

Table 1. Main LDACS1 OFDM parameters.

3.1 Frame structure

The 64 subcarriers in the FFT bandwidth are assigned to different types of symbols providing different functionalities:

- Null symbols are not transmitted. In most frame types, seven subcarriers on the left and six on the right hand side of the spectrum carry null symbols and serve as guard bands. In addition, the subcarrier in the center (DC subcarrier) of the spectrum is not transmitted.
- Pilot symbols are known in advance, exploited in the receiver for estimating the transmission channel.
- Symbols for reducing the Peak-to-Average Power Ratio (PAPR) are determined depending on the data in the respective OFDM symbol. PAPR reduction symbols are used in RL only.
- Synchronization symbols are used to obtain time and frequency synchronization in the receiver.
- Preamble symbols are used for facilitating receiver Automatic Gain Control (AGC).
- Data symbols are used for data transmission.

Multiple OFDM symbols are organized into frames. Depending on their functionality and on the link direction, different frame types are distinguished.

3.1.1 FL OFDM frame types

In the FL, BroadCast (BC) and combined Data/Common Control (CC) frames are utilized. The FL Data/CC frame comprises 50 subcarriers with 54 OFDM symbols, starting with two
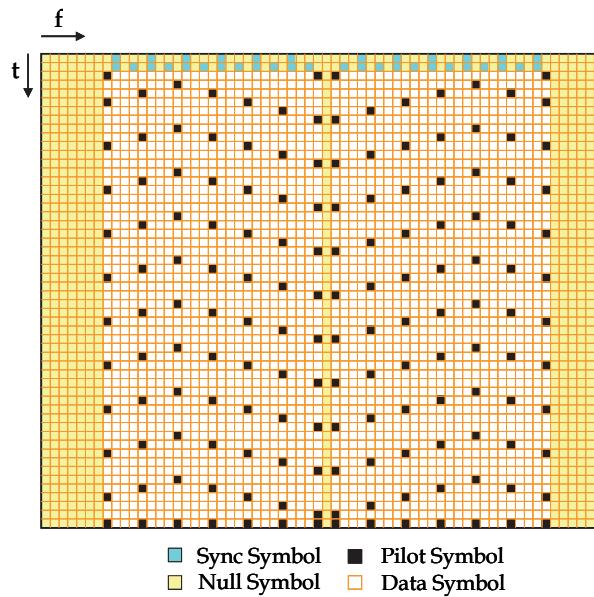
Fig. 3. Structure of a FL Data/CC frame.
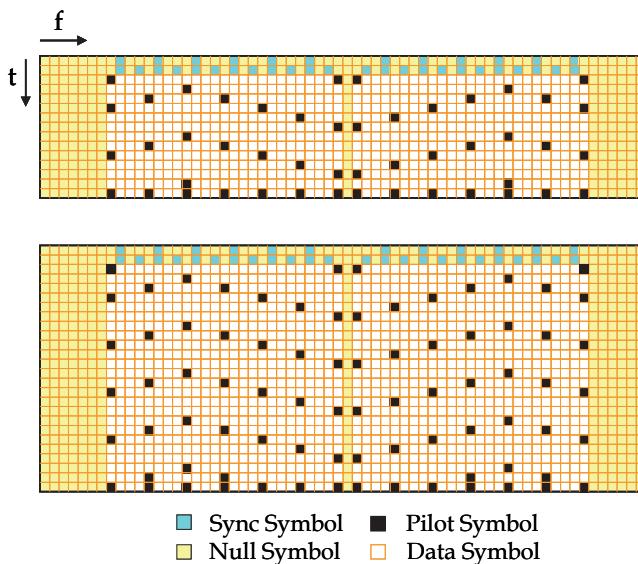
Fig. 4. Structure of BC1 and BC3 sub-frames (above) and BC2 sub-frame (below).

synchronization OFDM symbols followed by 52 OFDM symbols carrying data and pilot symbols, as depicted in Fig. 3. Subtracting the total number of 158 pilot symbols results in a total data capacity of 2442 symbols per FL Data/CC frame. The mapping of CC information and data onto this frame type is described later in this section.

An FL BC frame consists of three consecutive sub-frames (BC1/BC2/BC3), in which the Ground Station (GS) broadcasts signaling information to all active Airborne Stations (ASs) within its coverage range. Fig. 4 shows the structure of these sub-frames. In the BC1 and BC3 sub-frames, two of 15 OFDM symbols are used for synchronization. In the remaining 13 OFDM symbols, 48 carriers are modulated by pilot symbols and 602 remain for data transmission. The BC2 sub-frame is eleven OFDM symbols longer than the BC1/3 sub-frame and provides a capacity of 1120 data symbols.

Note that in the FL frames, no PAPR reduction symbols are inserted as the power amplifier used in the GS is assumed to provide sufficient linearity.

3.1.2 RL OFDM frame types

To realize multiple-access via OFDMA-TDMA in the RL, the transmission is organized in segments and tiles rather than in OFDM frames and sub-frames as in the FL. The usage of tiles enables the optimization of the resource assignments by the MAC sub-layer. Furthermore, bandwidth and duty cycle can be optimally selected according to the interference conditions.

As illustrated in Fig. 5, one tile, representing the smallest allocation block in the RL, spans 25 symbols in frequency and six symbols in time direction in the time-frequency plane. It comprises four PAPR reduction symbols and twelve pilot symbols. This leads to a data capacity of 134 data symbols per tile.
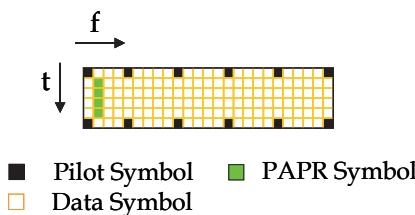

Fig. 5. Structure of a data tile in the RL.

The values of the PAPR reduction symbols are optimized in dependence on the data content such that the entire OFDM symbol, i.e. 24 data symbols and one PAPR reduction symbol, produces a minimal PAPR. In the first and the last OFDM symbol in a tile, PAPR is minimized by optimizing the phases of the pilot symbols in a way that the contribution of the pilot symbols to the PAPR is minimal.

The tiles of the data segment are subsequently positioned left of the DC subcarrier and mirrored versions of the tiles are positioned right of the DC subcarrier. The length of the data segment is kept variable by allocating a variable number of tiles to the data segment. This structure also supports a flexible assignment of resources to different ASs by assigning different tiles or different blocks of subsequent tiles to different ASs.

Signaling information like resource allocation requests is transmitted in the Dedicated Control (DC) segment. A DC segment has the same tile structure as the RL data segment.

However, it starts with a so called synchronization tile, spanning five OFDM symbols in time direction and 51 subcarriers, including the DC subcarrier, in frequency direction. The synchronization tile as illustrated in Fig. 6 starts with an AGC preamble, followed by two OFDM synchronization symbols. The forth and fifth OFDM symbol consist of pilot symbols only. This synchronization tile provides a possibility for an AS to execute seamless handover, when entering a new LDACS1 cell.
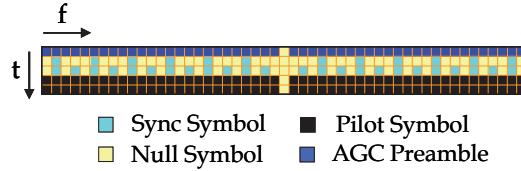

Fig. 6. Structure of one synchronization tile.

After the synchronization tile, an AGC preamble is inserted in the DC segment. This additional AGC preamble is necessary, since an AS, performing seamless handover is not yet power-controlled with the new GS, leading to a different power level at the receiving GS, compared to the signals from the other ASs. Within the remainder of the DC segment, exactly one tile is assigned to each AS within the cell. The preceding AGC preamble is transmitted by the same AS which transmits the first tile in the DC segment. The length of a DC segment is variable, depending on the number of ASs within the LDACS1 cell. As an example, one DC segment comprising six tiles corresponding to six active ASs within one cell, is depicted in Fig. 7.
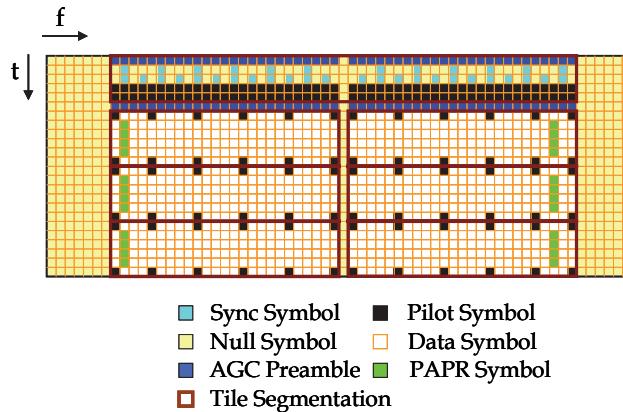

Fig. 7. Structure of one DC segment.

Besides the mentioned seamless handover, LDACS1 provides additional opportunities for handover by means of two consecutive RL Random Access (RA) opportunities, in which an AS can send its cell entry request to the GS (see Fig. 8). Propagation guard times of up to 1.26 ms precede and follow each RA frame. The propagation guard time of 1.26 ms corresponds to a maximal AS-GS distance of 200 nm. When transmitting a cell entry request, an AS is not yet synchronized to the GS, which means that the distance between the AS and

the GS is unknown. Hence, the distance can take values up to 200 nm which corresponds to the maximum LDACS1 cell size. The unknown propagation time differences are well compensated by the inserted guard times, guaranteeing that the cell entry request does not superimpose at the receiving GS with signals from other ASs within the cell.
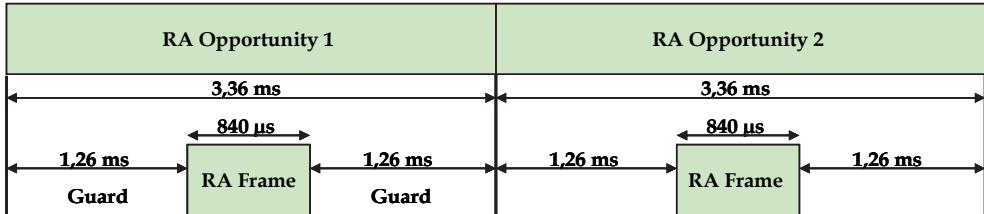

Fig. 8. Random access opportunities.

In Fig. 9, the structure of a RA frame itself is depicted. The first OFDM symbol represents the AGC preamble, the following two OFDM symbols contain synchronization sequences, while the remaining four OFDM symbols carry data, PAPR reduction symbols and pilot symbols. These four OFDM symbols use only 27 subcarriers (including the DC subcarrier), which leads to larger guard bands with 19 empty subcarriers on the left side and 18 on the right side.
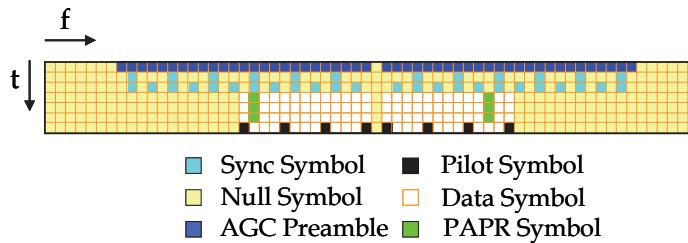

Fig. 9. Structure of random access frame.

3.1.3 Framing structure

In the FL, nine CC/Data frames are combined to one Multi-Frame (MF). The assignment of the nine frames within one MF either for user data or for common control information is variable. Starting with the fifth frame, one up to four frames can be allocated for common control information. The remaining frames contain user data. The mapping of modulated symbols onto the frames is performed block wise. These blocks are called Physical layer Protocol Data Units (PHY-PDUs). In general, three PHY-PDUs are mapped onto one Data/CC frame in FL MFs.

Each MF in the RL starts with a RL DC segment, followed by a RL data segment. The size of the DC segment, and thus also the size of the data segment is variable. The minimum size DC segment comprises one synchronization tile and the subsequent AGC preamble, followed by two tiles, corresponding to the case of one or two ASs within the cell. Since the extension of the DC segment over the entire MF is not reasonable, the maximum size of the DC segment is limited to 52 tiles. The remainder of the frame is filled up with the data segment. In the RL, one PHY-PDU is always mapped onto one tile.

The MF structure for FL and RL is shown in Fig. 10. The reference synchronization point for the FL and RL is the beginning of the MF. It is noticeable that the common control information in the FL and the dedicated control information in the RL are transmitted interleaved rather than simultaneously. The temporal shift of the FL control information allows the requests sent in the RL DC segment to be answered already in the FL CC frames of the same MF. Similarly, resource allocations transmitted in the FL CC frames can already be used in the next RL MF.
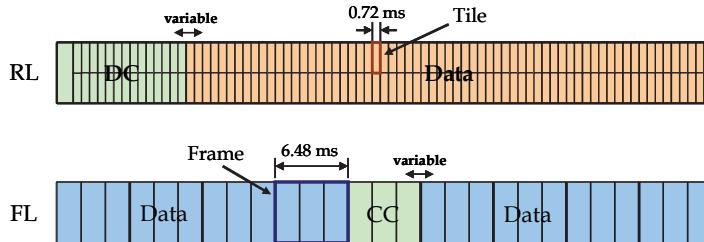

Fig. 10. Multi-Frame structures.

On top of the MF structure, a Super-Frame (SF) structure is provided. In the FL, one SF contains a BC frame of duration 6.72 ms, and four MFs, each of duration 58.32. In the RL, each SF starts with two opportunities for transmitting RL RA frames followed by four MFs. The start of the FL BC frame is synchronized to the start of the RL RA. The SF structure is summarized graphically in Fig. 11.
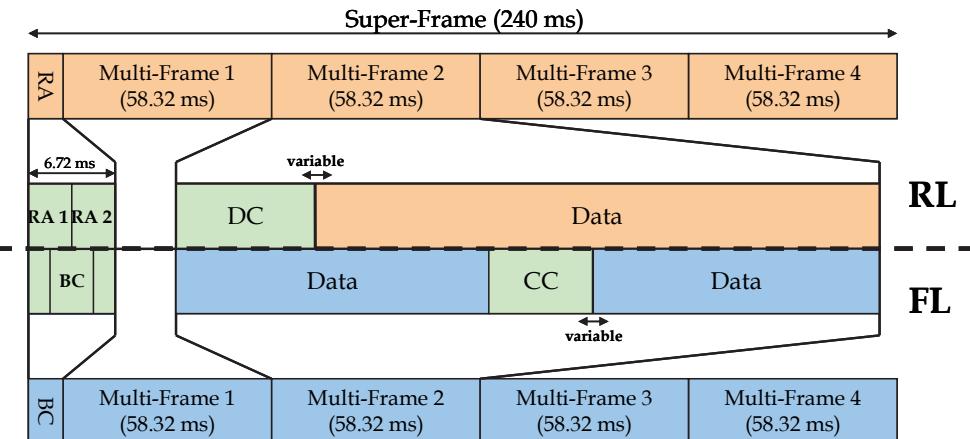
Fig. 11. Super-Frame Structure.

3.2 Coding and modulation

Different robust coding schemes proposed for LDACS1 should cope with the interference the transmit signal is exposed to. Primarily for small coding block sizes, a convolutional code of rate 1/3, followed by a bit interleaver is defined. For larger coding block sizes, a concatenated coding scheme consisting of a rate 0.9 Reed-Solomon (RS) code and a rate 1/2

convolutional code is employed. Again, the convolutional coder is followed by a bit interleaver. In addition, a byte interleaver between the two coders is inserted. The rate of the convolutional code can be changed from 1/2 to 2/3 or 3/4 by puncturing. The coding scheme is depicted in Fig. 12.

```
┌─────────┐    ┌─────────────┐    ┌─────────────┐    ┌─────────────┐
│RS Coder │───▶│   Byte      │───▶│ Convolutio  │───▶│    Bit      │───▶
│         │    │ Interleaver │    │ nal Coder   │    │ Interleaver │
└─────────┘    └─────────────┘    └─────────────┘    └─────────────┘
```
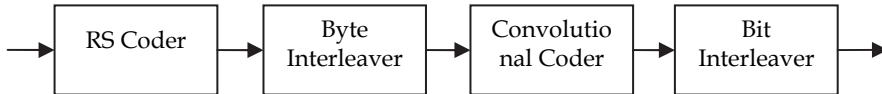
Fig. 12. Concatenated coding scheme.

For the subsequent modulation, Quaternary Phase-Shift Keying (QPSK) modulation, 16 Quadrature Amplitude Modulation (QAM), and 64 QAM are available.

The coding rates and the modulation scheme in the data frames can be adapted, based on the current channel and interference conditions, maximizing the transmission capacity. Therefore two Adaptive Coding and Modulation (ACM) modes are defined:

- Cell-specific ACM mode, which means that data for all users within one cell are encoded and modulated with a fixed scheme, and
- User-specific ACM mode, which means that separate coding and modulation schemes are applied for the data of different users.

In both modes, QPSK, 16QAM, and 64QAM are available. The rate of the convolutional code can be changed from 1/2 to 2/3 or 3/4. From the possible 9 combinations 8 are available[2] as ACM schemes.

For the particular frame types, the following coding and modulation scheme is chosen:

- In FL data frames, the concatenated coding scheme is mandatory. The coding rate and the modulation scheme is variable, both ACM modes are available. In case of cell-specific ACM, the information about the chosen coding and modulation scheme is transmitted in the BC frame. In case of user-specific ACM, the GS transmits the information about coding and modulation for different ASs via the coding and modulation scheme FL map in the CC information block.
- In RL data segments, the concatenated coding scheme is mandatory. The coding rate and the modulation scheme is variable, user-specific ACM is available. The selection of a coding and modulation scheme for a certain AS is carried out by the GS and communicated when assigning resources to this AS.
- For the BC sub-frames and the CC frames, again the concatenated coding scheme is employed. To guarantee an adequate protection of the control information, convolutional coding rate 1/2 and QPSK modulation is mandatory.
- The data in the RA frames and DC segment is encoded with the rate 1/3 convolutional coder, which is suitable for the small block sizes in these frames. QPSK modulation scheme is mandatory.

3.3 Reduction of out-of-band radiation

High out-of-band OFDM side lobes may cause harmful interference at the Rx of other L-band systems and have to be reduced. For that purpose, Tx windowing is applied in order to smooth the sharp phase transitions between consecutive OFDM symbols which cause out-of-band radiation.

[2]Only the combination of 16QAM with rate 3/4 coding has been omitted.
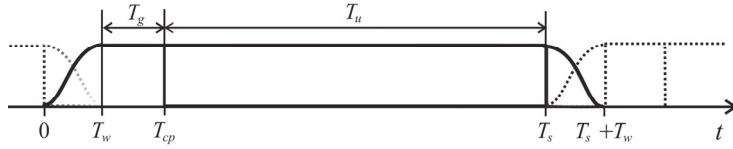
Fig. 13. Tx windowing principle.

As illustrated in Fig. 13, the OFDM symbol is extended by a cyclic prefix of total length $T_{cp}=T_w+T_g$. One part of the cyclic prefix serves as guard interval of length T_g to compensate multipath propagation on the radio channel. The second part is of length T_w and contains the leading edge of the window. In addition, the OFDM symbol is extended by a cyclic suffix which contains the trailing edge of the window of length T_w. This approach guarantees that Tx windowing does not affect the useful part of the OFDM symbol including guard interval. To keep the overhead induced by extending the OFDM symbol duration at a minimum, subsequent OFDM symbols overlap in those parts containing the leading and trailing edges of the window. For LDACS1, a raised-cosine window with roll-off factor $\alpha = 0.107$ is proposed.

3.4 Receiver design
When deployed as an inlay system, signals of existing L-band systems may cause severe interference onto the LDACS1 Rx. Especially DME systems operating at a small frequency offset to the LDACS1 channel represent a source of strong interference. Two promising techniques, aiming at mitigation of the influence of interference onto LDACS1, are presented in the following. In addition, the effect of the interference and the proposed interference mitigation to the estimation of the transmission channel is examined.

3.4.1 Over-sampling
RF and IF filters in the selective stages of the LDACS1 Rx successively reduce interference contributions received outside the LDACS1 bandwidth. Since the interference signal power can be very high, it may be impossible to completely remove this out-of-band interference power. In particular, due to the relatively large occupied bandwidth of the DME signal, the spectra of DME Txs operating at ±0.5 MHz offset from the LDACS1 channel partly fall into the LDASC1 Rx bandwidth. Subsequent sampling of the filtered Rx signal in the A/D-converter with the native sampling period leads to periodic repetitions of the interference spectra in the frequency domain in distances of multiples of the FFT bandwidth B_{FFT}. In the case without interference, the filtered Rx signal is band-limited by B_{FFT}. Sampling with $T_{sa} = 1/B_{FFT}$ does not lead to aliasing effects. However, the Nyquist sampling theorem is not fulfilled for the interference signal. The remaining out-of-band interference signal is not band-limited to B_{eff}, hence due to aliasing after Rx FFT operation the undesired signal parts fall into the LDACS1 bandwidth. This can be avoided by over-sampling the time domain Rx signal at least by a factor of four, resulting in an increased spacing between the periodic repetitions in the frequency domain and reduced aliasing effects. The over-sampled Rx signal is then transformed to the frequency domain by FFT with the size increased according to the over-sampling factor. For further signal processing, only the relevant subcarriers in the LDACS1 system bandwidth are considered. All other subcarriers are discarded.

3.4.2 Pulse blanking

Pulse Blanking (PB) is a well-known approach for reducing pulsed interference such as DME interference. It has already been applied for reducing DME interference in the E5- and L5-bands used by satellite navigation systems (Gao, 2007) and for reducing impulsive noise in OFDM systems (Zhidkov, 2006).

In LDACS1, PB is applied to the digitized Rx signal after A/D conversion. When the amplitude of the over-sampled Rx signal exceeds a pre-defined threshold T_{PB} the corresponding samples are blanked, i.e. set to zero. Besides the DME interference, the blanked samples also comprise the desired OFDM signal and noise. Hence, the threshold T_{PB} has to be chosen carefully. When choosing it too low, the corruption of the desired signal exceeds the benefit of reducing the interference power, whereas a too high threshold leads to a strong remaining interference power. A suitable criterion for optimally setting T_{PB} is the Signal-to- Interference-and-Noise ratio (SINR) (Zhidkov, 2006).

Due to high PAPR in OFDM systems, peaks of the desired OFDM signal and DME pulses cannot be clearly distinguished. For an unambiguous detection of interference pulses in the Rx signal, a correlation of the Rx signal with a known interference pulse has been proposed in (Brandes & Schnell, 2009).

An algorithm for diminishing the harmful PB influence onto the useful OFDM signal was presented in (Brandes et al., 2009). It is based on the idea, that the PB impact on the OFDM signal can be determined exactly when representing PB as a windowing operation. The window function is a rectangular window that exhibits notches at those positions where the Rx signal is blanked. Recalling that the shape of the window determines the spectrum of the OFDM subcarriers, the subcarrier spectra can be determined and the distortion induced by PB is identified as Inter-Carrier Interference (ICI). ICI can easily be reduced by subtracting the known impact of all other subcarriers from the considered subcarrier as applied for example for reducing ICI in OFDMA systems induced by frequency offsets (Fantacci et al., 2004).

3.4.3 Channel estimation

As basic channel estimation algorithm in LDACS1, a pilot-aided linear interpolation is proposed. In order to make channel estimation robust towards interference the pilot tones in frequency direction are spread over all subcarriers in order to reduce the number of pilot tones that would be affected by a strong DME interference pulse from the adjacent DME channel. In RL, such pilot tone placing is not possible due to the OFDMA approach. In this case the pilot tones have to be placed of the edges of each tile. These pilot distances in time and frequency direction have been chosen in accordance with the expected Doppler shifts and maximal delay times in the case of multipath propagation. To improve channel estimation, pilot boosting is applied. In this case, the power of the pilot symbols is increased by 4 dB over the average power of each data symbol.

A more sophisticated approach for estimating the transmission channel is Wiener filtering. The coefficients of the Wiener interpolation filter are derived by minimizing the Mean-Squared-Error (MSE) between the actual and the estimated channel coefficients. This leads to an optimal noise suppression, given the noise variance and channel statistics. In (Epple & Schnell, 2010), it was proposed to incorporate an equivalent noise term, comprising the additive white Gaussian noise and the estimated DME interference, into the Wiener filter. If interference mitigation like PB is applied, the variance of the induced ICI can be

incorporated into the Wiener filtering (Epple et al., 2011). Both approaches turned out to improve the channel estimation performance remarkably.

4. Conclusion

LDACS1 is the broadband candidate for the future L-band communications system which covers the air-ground link within the FCI. The design of LDACS1 is based on the multi-carrier technology OFDM, a modern communications technology which is highly flexible and efficient. In comparable application domains, like mobile radio communications, OFDM is the current state-of-the-art solution and is also foreseen for the next generation of mobile radio systems.

The LDACS1 design is based on existing multi-carrier standards, like WiMAX and P34. However, due to the operational requirements on an aeronautical communications system, the propagation conditions, and the interference environment in L-band a specific OFDM solution had to be designed. Investigations of the LDACS1 system performed by computer simulations have proven the suitability of the LDACS1 design for use in the L-band (Brandes et al., 2009; Epple & Schnell, 2010; Epple et al., 2011). Even for the challenging inlay deployment scenario, LDACS1 is expected to work according to the requirements without causing harmful interference to the legacy L-band systems. With that, LDACS1 shows that aeronautical communications can profit from the developments in related fields and can achieve efficient usage of the scarce spectrum resources currently available for communications within aviation.

5. References

Bello P.A. (1973). Aeronautical Channel Characterization, *IEEE Trans. on Communications*, vol. COM-21, no 5, pp. 548-563, May 1973

Brandes, S.; Epple, U. & Schnell, M. (2009). Compensation of the Impact of Interference Mitigation by Pulse Blanking in OFDM Systems, *Proceedings of IEEE Global Telecommunications Conference 2009*, Honolulu, Hawaii, USA, November 28-December 4, 2009

Brandes, S. & Schnell, M. (2009). Interference Mitigation for the Future Aeronautical L-Band Communication System, *Proceedings of 7th International Workshop on Multi-Carrier Systems & Solutions 2009*, Herrsching, Germany, May 5-6, 2009

Budinger J.M. (2011). Aeronautical Mobile Airport Communications System (AeroMACS), Future Aeronautical Communications, Simon Plass (ed.), ISBN 979-953-307-443-5

Epple, U.; Shibli, K. & Schnell, M. (2011). Investigation of Blanking Nonlinearity in OFDM Systems, *Proceedings of IEEE International Communications Conference 2011*, Kyoto, Japan, June 5-9, 2011

Epple, U. & Schnell, M. (2010). Channel Estimation in OFDM Systems with Strong Interference, *Proceedings of 15th International OFDM Workshop 2010*, pp. 26-30, Hamburg, Germany, September 1-2, 2010

EUROCONTROL (2007). Future Communications Infrastructure - Technology Investigations. Description of AMACS, v1.0, 2007. Available from http://www.eurocontrol.int/communications/public/standard_page/LBANDLIB .html. Information available also on http://www.eurocontrol.int

EUROCONTROL & FAA (2007). Communications Operating Concept and Requirements for the Future Radio System, (COCR) version 2

Fantacci, R.; Marabissi, D. & Papini, S. (2004). Multiuser Interference Cancellation Receivers for OFDMA Uplink Communications with Carrier Frequency Offset, *Proceedings of IEEE Global Telecommunications Conference 2004*, pp. 7081-7085, Dallas, Texas, USA, November 29-December 3, 2004

Fistas N. (2011). Aeronautical Future Aeronautical Communications: The Data Link Component, Future Aeronautical Communications, Simon Plass (ed.), ISBN 979-953-307-443-5

Gao, G.X. (2007). DME/TACAN Interference and its Mitigation in L5/E5 Bands, *Proceedings of ION Institute of Navigation Global Navigation Satellite Systems Conference*, Fort Worth, Texas, USA, September 25-28, 2007

Gligorevic, S., Zierhut, R., Jost, T., Wang, W. (2009). Airport Channel Measurements at 5.2GHz, *In Proceedings of the 3rd European Conference on Antennas and Propagation (Eucap),* March 2009

Gräupel T. & Ehammer M. (2011). The LDACS1 Link Layer Design, Future Aeronautical Communications, Simon Plass (ed.), ISBN 979-953-307-443-5

Haindl, B.; Rihacek, CHr.; Sajatovic, M.; Phillips, B.; Budinger, J.; Schnell, M.; Lamiano, D. & Wilson, W. (2009). Improvement of L-DACS1 Design by Combining B-AMC with P34 and WiMAX Technologies, Integrated Communications Navigation and Surveillance Conference (ICNS 2009), Arlington, VA, USA, May 2009

ICAO, ACP-WGC (2006). Report of the Tenth Meeting of ICAO ACP Working Group C, Montreal, 13-17 March 2006

ICAO, ACP-WGC (2011). FCS Phase II Results Paper 3 - Detailed Technology Investigations, Working Group C – 11th meeting, Brussels, Belgium, 18 – 20 September 2006

IICAO, ACP-WGW (2008). Report of the Working Group of the Whole, Second Meeting, Montreal, 21 to 25 April 2008

Matolak D.W., Sen I., Xiong W. (2008). The 5GHz Airport Surface Area Channel Part I, Measurement and Modeling Results for Large Airports, *IEEE Trans. Vehicular Tech.*, vol. 57, no. 4, pp. 2014–2026, 2008

Rice, M.; Davis, A.; Bettweiser, C. (2004). Wideband Channel Model for Aeronautical Telemetry; *IEEE Trans. On Aerospace and Electronic Systems*, vol. 40, no.1, January 2004

SESAR JU (2011). Updated LDACS1 System Specification. Information available at www.sesarju.eu

Zhidkov, S.V. (2006). Performance Analysis and Optimization of OFDM Receiver with Blanking Nonlinearity in Impulsive Noise Environment, *IEEE Transactions on Vehicular Technology,* Vol. 55, No. 1, (January 2006), pp. 234–242

Part 5

Visions for Aeronautics

IFAR – The International Forum for Aviation Research

Richard Degenhardt, Joachim Szodruch and Simon Plass
German Aerospace Center (DLR)
Brunswick, Cologne and Oberpfaffenhofen,
Germany

1. Introduction

The future challenges of air transport motivated the leading worldwide aviation research institutions to found IFAR - the International Forum for Aviation Research which aims at discussing the global aeronautical challenges and the set-up of a Framework outlining worldwide research. Climate change is currently the most relevant topic and was the motivation to set up IFAR. However, IFAR also addresses other topics relevant for a global air transport system (e.g. noise, security, safety, efficient operations). IFAR connects and represents worldwide aviation research and provides a common voice for their members in the international dialogue. IFAR interacts with the society, global politics and industry and takes up the challenges identified by them.

The idea of IFAR was born at the Berlin Summit 2008 where key leaders of 12 international aeronautical research organisations met to address the question of future Air Transport in the context of climate change. In this regard, the participants agreed that any research and strategy contributing to new solutions will have to reconcile the increasing need for international mobility in a globalized work-sharing economy with the challenge of simultaneously developing new solutions to balance the climate effects of the accompanying world-wide air traffic growth. At the second Berlin Summit in 2010 16 international aeronautical research organisations met and eventually set up IFAR (IFAR, 2008).

The main objective of IFAR is connecting global research establishments and setting up a Framework agreed upon by research institutions worldwide. Within this document promising technologies will be identified which contribute to an improved Air Transport System. IFAR as research representative focuses on the identification of new technologies up to the development of Technology Readiness Level (TRL) level 6. The IFAR members agreed at the first IFAR Summit 2010 to focus in 2010/2011 on the topics related to climate change and to present potential solutions for an ecologically and economically efficient air transport system. Within the next years this Framework is going to be extended by taking the other topics noise, safety, security and efficient operations into account (Szodruch et al., 2011a).

This paper deals with the objectives, state-of-the art and future planning of IFAR. It highlights first ideas for improved technologies in the area Aeronautical Communications which is the main topic of this book. Aeronautical Communications is one relevant topic considered in IFAR which plans to contribute to an improved air transport system on a

worldwide level. A communications network is to be created which meets the requirements of the aviation industry of the future. Here, the data streams must, above all, flow reliably between the aircraft and the ground, and this must take place both in remote regions over the oceans and the poles as well as in crowded conurbations. Supplementary information means that such a new communications network can sustainably improve safety standards in aviation and also reduce environmental impact through optimised flight paths for example. Within Section 6 the IFAR future aeronautical communications aspects will be spotlighted.

2. IFAR history

The Forth Assessment Report of the International Panel on Climate Change (IPCC) has stirred an intensive public debate on future aeronautical research challenges and policies. By an initiative of the German Aerospace Center (DLR) a Summit was held in 2008 in Berlin as a response. 12 key international leaders in aeronautical research met to address the question of the Air Transport of the Future in the context of climate change. In this regard, the conference participants agreed that any research and strategy contributing to new solutions will have to reconcile the increasing need for international mobility in a globalized, work-sharing economy with the challenge of simultaneously developing new solutions to balance the climate effects of the accompanying world-wide growth in air traffic. The IPCC report identifies aviation to contribute 2–3 percent of today's total global anthropogenic CO_2 emissions. This prompted the International Air Transport Association (IATA) to set the long term challenge of Zero Emission Aviation by 2050 and emphasised the importance of addressing these challenges on a global level. Using the IPCC report and the latest research results on climate change as a basis, the Berlin Summit participants acknowledged the need for new solutions addressing both mid-term as well as long-term perspectives. Some solutions, such as the enhanced efficiency of aircraft and air traffic management systems have already resulted in major technological advancements and increased operational capabilities. Nonetheless, as air transport faces increasing demands, these topics remain to be core areas of aeronautical research. Accordingly, international research establishments are key actors, particularly in approaching the long-term and, thus, pre-industrial questions of research and development. The participants of the Berlin Summit welcomed this first event as a unique international forum to enhance discussion of the strategic challenges in aeronautical research. They agreed to establish an international platform for a dialogue to coincide with International Air Shows and with meetings all over the world.

At the Berlin Summit in 2010 key leaders of 16 international aeronautical research organisations met the second time and gave this forum the name IFAR which stands for International Forum for Aviation Research. The attendees continued the discussion on the topics related to climate change and agreed to develop a common Research Framework which represents the Aviation Research worldwide. The kind of organisation is under discussion and will be defined in the short future.

The outcome of the IFAR summit was summarised by the participants with in a declaration which is published at the IFAR website www.ifar.aero.

3. IFAR objectives

The objectives for IFAR were discussed at the last IFAR Summit in 2010 and the outcome was published within a declaration. These results are at this time first ideas which will be finalised in the short future. This section gives a summary of this declaration.

IFAR is a forum which represents the aviation research organisations worldwide. The specifics of the organisation (e.g. alliance) are currently under discussion and will be defined in the near future. Fig. 1 illustrates the interaction of IFAR with the society, global political influence organisations and industry. IFAR takes up the challenges identified by them and develops as reaction research strategies.
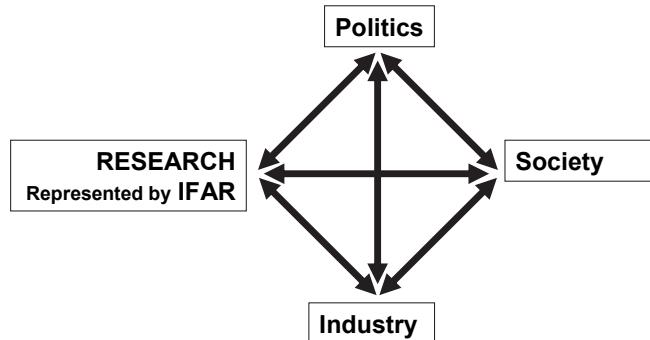

Fig. 1. Interaction of IFAR to politics, public and industry.

The main IFAR aims are
- to take up the environmental challenges for air transport industry as identified by the global community
- to present potential solutions for an ecologically and economically efficient air transport system
- to connect worldwide aviation research institutions
- to provide a common voice for the IFAR organisations in the international dialogue
- to focus IFAR's initiatives on the technologies to address climate change, impact of weather and natural phenomena, noise and local emissions, efficient operations, security and safety
- to concentrate at the beginning of IFAR's initiatives on climate change
- to develop an IFAR Framework considering other international aviation road maps and research initiatives.

IFAR is considering for the development of the IFAR Framework the following technological and other aviation topics which are of global nature:
- Technological topics
  - Climate change
  - Alternative fuels
  - Noise
  - Security
  - Safety and
  - Efficient operations
- Other topics
  - Education
  - Quality
  - Crisis management
  - Networking
  - Capacity

Concerning the development of the Framework the participants agreed to the plan illustrated in Fig. 2. The partners started on the topic climate change which is mostly discussed in public and will continue in the next years with the other topics as noise, security, safety and efficient operations. IFAR will disseminate its work and progress to the public regularly at www.ifar.aero.
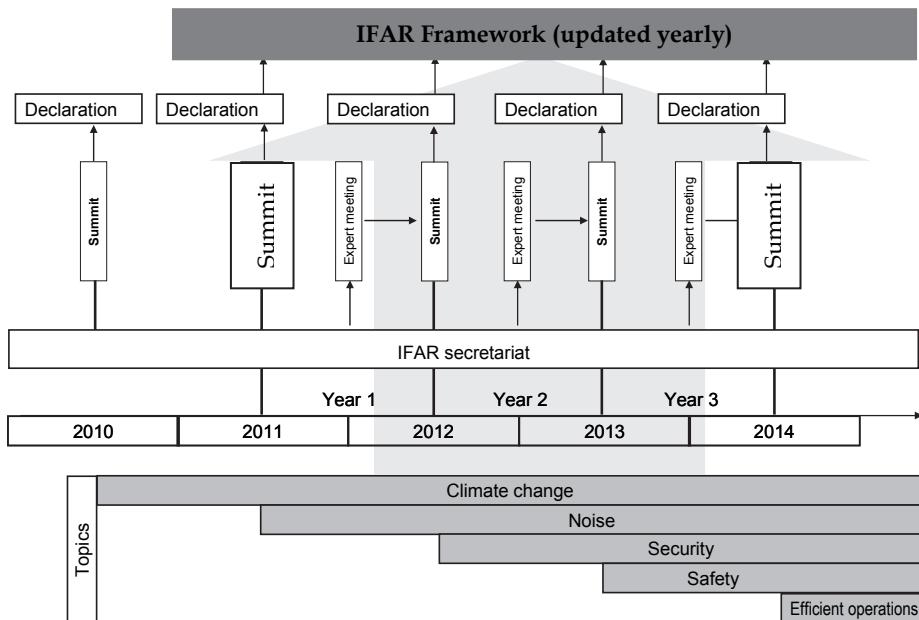
Fig. 2. IFAR meetings related to topics.

4. Aviation research - state of the art

4.1 History

Over the past 100 years aviation has transformed the society dramatically. Looking back at the last 50 years the aviation passed a spectacular development. The International Energy Agency (IEA) developed the graph in Fig. 3 which shows for that time the improvement of the energy intensity (fuel burn per passenger kilometre) for selected aircraft. This figure illustrates that the technology in engine, airframe and other measures has helped to reduce the aircraft fuel burn per passenger kilometre by more than 70%. This is already an excellent success. However, a significant growth of the Air Traffic System (cf. next section) is expected in the next years. Due to the negative impact on the climate and the decreasing availability of fuel resources there is still a high demand for a further improvement of the energy intensity. It is the responsibility of the aviation research to develop the corresponding new technologies as well as looking into alternative fuels.

4.2 Outlook into the future
4.2.1 General CO_2 forecast

IEA published in 2010 the Energy Technology Perspectives - Scenarios and strategies to 2050. This report (ETP, 2010) analyses and compares various scenarios. It does not aim to forecast what will happen, but rather to demonstrate the many opportunities to create a more secure and sustainable energy future. A comparison of different scenarios demonstrates that low-carbon technologies can deliver a dramatically different future. However, it is mandatory not only to stimulate the evolutionary development of new
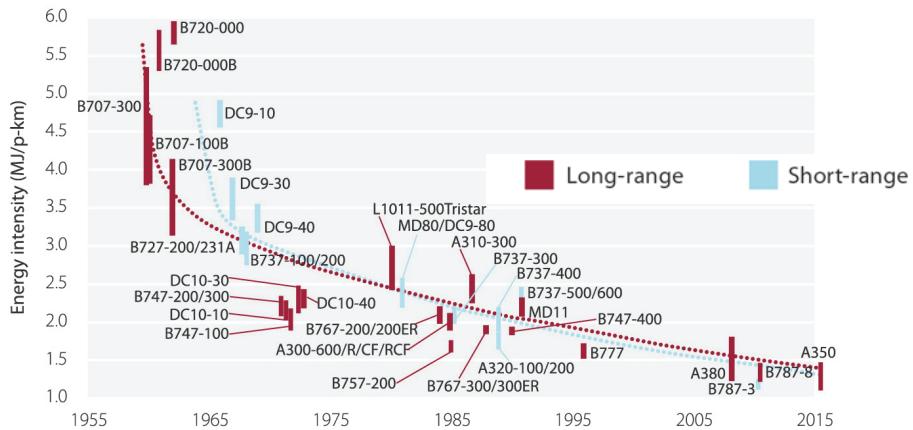
Fig. 3. Energy intensity of aircraft. The range of points for each aircraft reflects varying configurations; connected dots show estimated trends for short and long-range aircraft. (Source: IEA).

application oriented technologies but also to invest in revolutionary ideas and motivate creativity and fundamental research. Thus, simply increasing funding will not be sufficient to deliver the necessary low-carbon technologies. Current government RD&D programmes and policies need to be improved by adopting best practices in design and implementation. This includes:

- the design of strategic programmes to fit national policy priorities and resource availability;
- the rigorous evaluation of results and adjusting support if needed;
- and strengthening the linkages between government and industry, and between the basic science and applied energy research communities to accelerate innovation

Current energy and CO_2 trends run directly counter to the repeated warnings sent by the United Nations Intergovernmental Panel on Climate Change (IPCC), which concludes that reductions of at least 50% in global CO_2 emissions compared to 2000 levels will need to be achieved by 2050 to limit the long-term global average temperature rise to between 2.0°C and 2.4°C. Recent studies suggest that climate change is occurring even faster than previously expected and that even the "50% by 2050" goal may be inadequate to prevent dangerous climate change (cf. Fig. 4 and Fig. 5).

Temperature increase (°C)	All GHGs (ppm CO_2-eq.)	CO_2 (ppm CO_2)	CO_2 emissions 2050 (% of 2000 emissions)
2.0-2.4	445-490	350-400	−85 to −50
2.4-2.8	490-535	400-440	−60 to −30
2.8-3.2	535-590	440-485	−30 to +5
3.2-4.0	590-710	485-570	+10 to +60
4.0-4.9	710-885	570-660	+25 to +85
4.9-6.1	885-1 130	660-790	+90 to +140

ppm: parts per million.

Fig. 4. Relationship between CO_2 emissions and climate change (ETP, 2010).
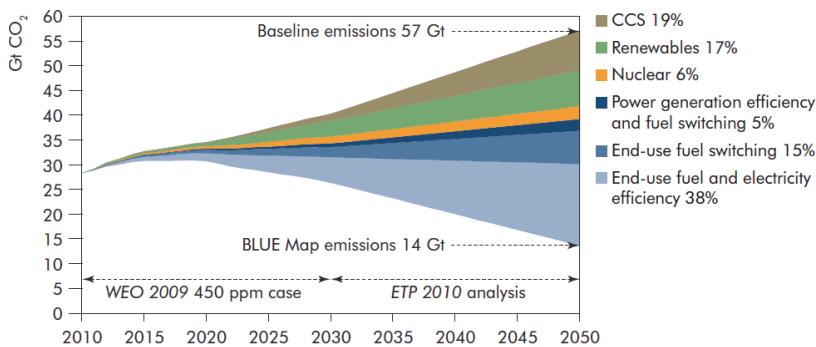
Fig. 5. Contribution of different technologies to CO_2 emissions (ETP, 2010).

4.2.2 Aviation

The current aviation's contribution to global CO_2 emissions is estimated at 2% and its contribution to total greenhouse gas emissions is approximately 3%, since other exhaust gases and contrails emitted during flight also contribute to the greenhouse effect. The aviation industry contributes approximately 8% to the world gross domestic product, and aviation growth is projected to be 5 to 6% per year (IATA (2009)). By 2050, the IPCC forecasts aviation's share of global carbon emissions will grow to 3% and its contribution to total greenhouse gas emissions is estimated to 5%.

According to (ETP, 2010) air travel is expected to be the fastest growing transport mode in the future as it has tended to grow even faster than incomes during normal economic cycles. Air passenger-kilometres increase by a factor of four between 2005 and 2050 in the Baseline scenario (no actions e.g. due to improved technologies, cf. Fig. 5) , or even by a factor of five in a High Baseline scenario. In the same period, aviation benefits from steady efficiency improvements in successive generations of aircraft. The technical potential to reduce the energy intensity of new aircraft has been estimated in a range between 25% and 50% by 2050. This is equivalent to an improvement of about 0.5% to 1% per year on average. Additionally, airlines show an improvement roughly by 2% in 10 years.

Fig. 6 and Fig. 7 depict the long-term growth of aviation, measured by revenue passenger kilometres and CO_2 emissions under different scenarios (Szodruch et al., 2011b):

- Scenario 1 represents the ATS up to 2050 with aircraft technology that is currently available. Improvements in fuel efficiency are therefore limited to the replacement of legacy aircraft currently operated with state-of-the-art technology.
- In scenario 2, a 50% reduction in specific fuel consumption (ACARE objective) is achieved by a combination of aircraft entering service after 2020, operational measures and improvements in air traffic management.
- Scenario 3 depicts a situation where CO_2 emissions are stabilised after 2030, without constraining aviation growth. This scenario requires considerable technological efforts in excess of the objectives, to achieve a stabilisation of emissions. In addition to operational improvements of the air transport system, the fuel efficiency of new aircraft types entering service after 2020 is required to increase by about 60% compared to the technology level of 2000.

The forecast of passenger traffic is based on the predictions of Airbus, and Boeing, which publish forecasts for up to 20 years, the International Civil Aviation Organisation's

(ICAO)(2007) Outlook for 2025, and the results of CONSAVE 2050; a study that quantified long-term scenarios to 2050 (Berghof et al., 2005).
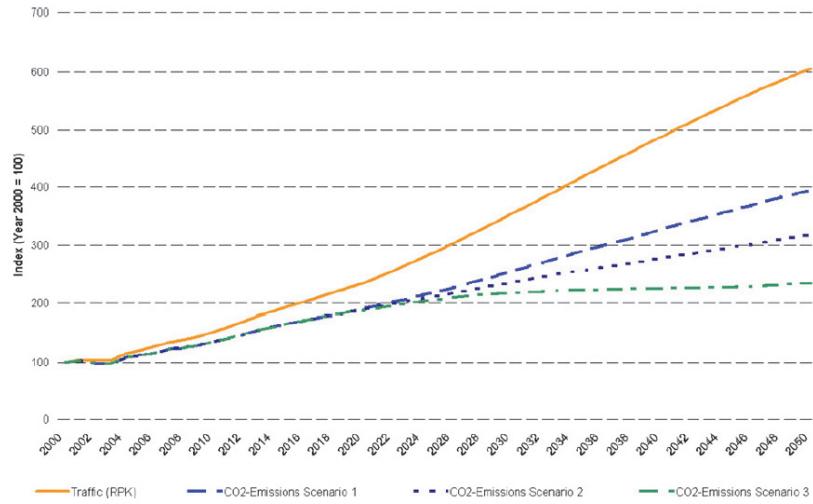

Fig. 6. Development of passenger traffic and CO_2 emissions 2000-2050.
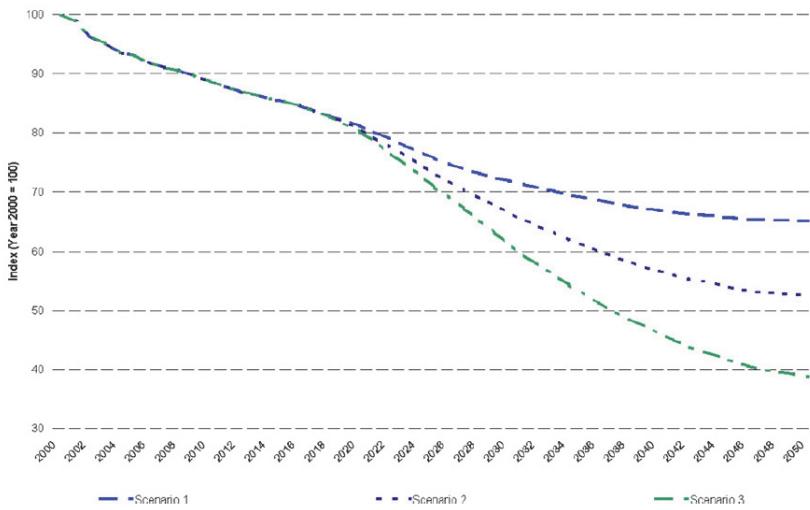
Fig. 7. Development of fleet-wide specific consumption 2000-2050.

5. IFAR Framework

5.1 IFAR approach

The IFAR approach consists of 3 steps illustrated in Fig. 8. Step 1 builds the IFAR vision 2050 which is mainly influenced by society, stakeholders and political demands (e.g. the

need for new technologies reducing influence on the climate). Step 2 considers new and visionary break trough technologies which are expected to fulfil the goals in Step 1 and to improve the Air Transport System (ATS) in Step 3. Technologies considered in this regard are not only software or hardware but also improved operations or other innovative ideas. IFAR - as research representative - concentrates on technologies until TRL 6. Further development, qualification and product integration can only be done by industry. The search for new technologies does not necessarily need to be conducted within the aviation sector. They can also be transferred from other industrial sectors as automotive, space, energy, etc. Alternative fuels, which might play an important role in the future ATS can for instance be developed in the energy sector. On the other hand the new technologies developed in aviation may also be transferred to other industrial fields. Aeronautics is for instance working on the automation of the manufacturing process for future aircraft structures made of composites. This technology may be partly transferred to other sectors different from aviation. Step 3 is the future Air Transport System improved by the new technologies from Step 2. The expected impact of single technologies or combinations of them on the ATS is also part of Step 2. The new ATS has to take the influence of numerous regulations into account.
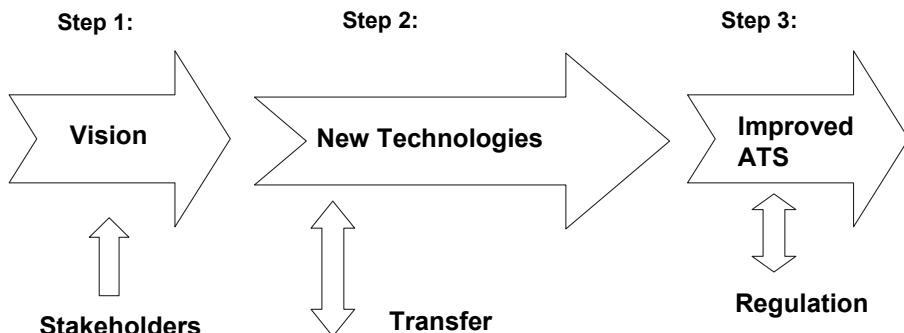

Fig. 8. IFAR approach.

The IFAR Framework is currently under development. It is planned to be a summary or harmonisation of available strategic documents provided by the IFAR partners. Two documents are public (from European Research Establishments in Aeronautics (EREA) which represents Europe (EREA, 2010) and from NASA (NASA, 2010) and other input is expected to be provided from IFAR discussions and further documentations by the partners. Strategic Road Maps of organisations outside IFAR will also be considered. Fig. 9 summarizes the public documents which contribute to the IFAR Framework, namely from the International Air Transport Association (IATA) (IATA, 2009), the International Energy Agency (IEA) (ETP, 2010), Advisory Council for Aeronautics Research in Europe (ACARE) (ACARE, 2010) or the Flightpath 2050 (Flightpath 2050, 2011).

Step 1: IFAR vision

Step 1 of the IFAR approach represents the IFAR vision which is influenced by stakeholders and by political demands. IFAR aims to develop an own target point in the vision for each single technological topic as climate change, noise, security, safety and efficient operations.
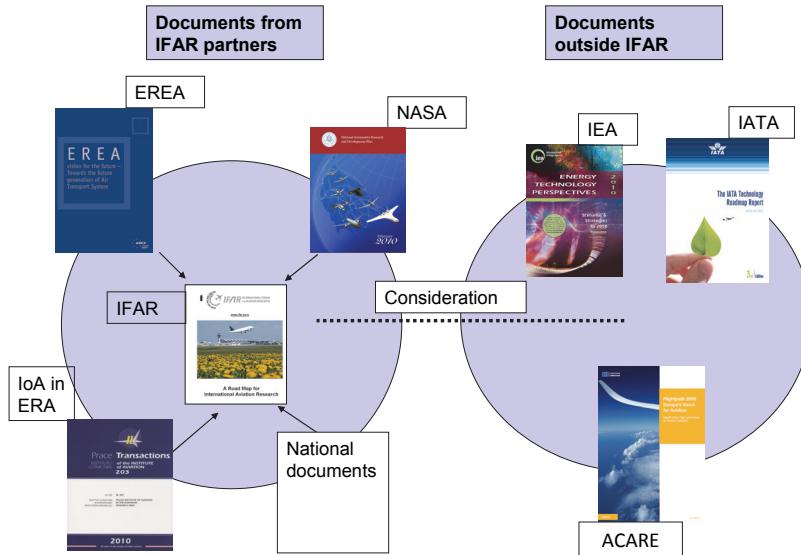
Fig. 9. Documents from IFAR partners and organisations outside IFAR.

For climate change there exist already for instance the following visions 2050 of IATA or IEA:

- IATA vision: 50% Reduction in net CO2 emissions over 2005 levels
- IEA vision for Aeronautics: ATS is operating with new energy sources by 30%.

IFAR is currently developing its own vision. For the topic climate change the already available visions from IATA or IEA will be taken into account, but the IFAR vision will be extended by the consideration of the total Air Transport System as well as the impact on the global temperature increase. Air transport impacts the climate directly for instance by contrails, soot, CO2, NOx and other emissions. All this leads to an increase of the global temperature. However, there are operational technologies (e.g. flying in different altitudes or routes) which have influence on the global temperature but not CO_2. Thus, the inclusion of the global temperature as an additional metric is reasonable and will allow a better evaluation of the impact of such technologies on the climate.

Step 2: New technologies

IFAR aims to identify promising and breakthrough technologies which are expected to fulfil the IFAR vision defined in Step 1. IFAR considers here for instance technologies improving the performance of the aircraft, the airport, the air traffic management (ATM), flights with low environmental impact (different altitudes or routing) or the interaction of all technologies together. Other examples are alternative fuels to reduce the carbon foot print of the Air Traffic System and minimise the independent of oil. The technologies considered in IFAR cover the full range of the ATS (cf. Fig. 10). The technologies are usually developed by the aviation sector itself but they may also be transferred from or to other industrial sectors as automotive, space or energy. IFAR is currently developing a technology tree which will be one main part of the IFAR Framework. The technologies will be the input from available IFAR documents provided by the IFAR partners (cf. Fig. 9).
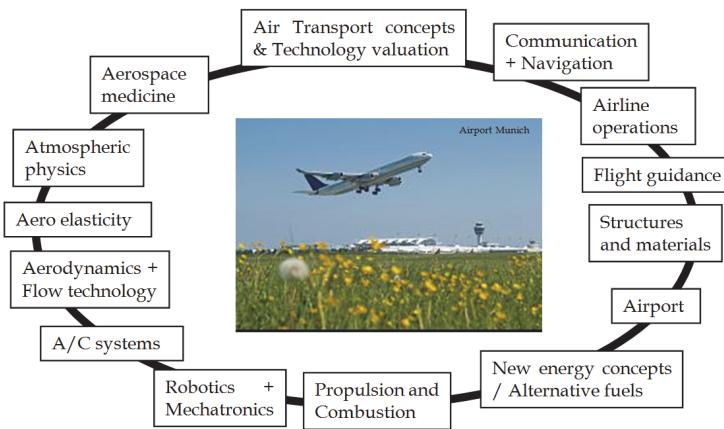
Fig. 10. Aviation topics considered in IFAR.

Step 3: Improved ATS

Step 3 of the IFAR approach represents the Future ATS. The improvement will be an outcome of the assessment of the new technologies discussed in Step 2. IFAR defines and agrees during expert meetings on the level of technology impact.

6. Communications aspects

Within the IFAR, communication and navigation are considered as an aviation topic. The technologies for the future communications infrastructure (FCI) are based on seamless networking and future data links. The concept of seamless networking describes the interoperability of all existing and future (digital) data links and service-oriented avionic architectures to allow a single infrastructure and information management system to deliver instantaneous data with high quality. To enable this concept, new data links with higher capacities, better flexibility, and increased coverage are needed. Fig. 11 shows a global aeronautical communication network.
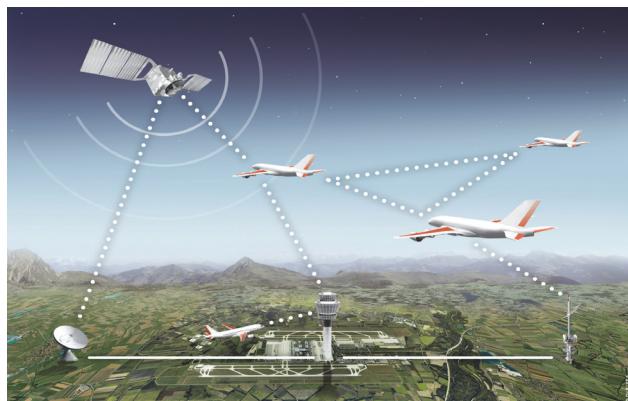

Fig. 11. Integration of different data links into a global aeronautical communication network.

6.1 Existing visions for ATM by 2020

The Single European Sky ATM Research Programme (SESAR) aims at developing the new generation ATM system capable of ensuring the safe and smooth air transport worldwide over the next 30 years. SESAR's goal to 2020 is saving 8 to 14 minutes, 300 to 500 kg of fuel and 948 to 1575 kg of CO_2 per average flight (SESAR, 2009).

The Next Generation Air Transportation System (NextGen) developed and planned to be implemented by the US Federal Aviation Administration (FAA) will allow more aircraft to safely fly closer together on more direct routes, reducing delays and providing unprecedented benefits for the environment and the economy through reductions in carbon emissions, fuel consumption and noise. By 2018, NextGen will reduce total flight delays by about 21 percent. In the process, more than 1.4 billion gallons of fuel will be saved during this period, cutting carbon dioxide emissions by nearly 14 million tons (NextGen, 2009).

One major pillar in the SESAR and NextGen concepts is the FCI to support the new operational concepts that are being developed.

The ACARE Vision beyond 2020 (and towards 2050) states a noise reduction by innovative mission and trajectory planning due to a better ATM. Furthermore, improved ATM and operational efficiency contribute by 5-10% to the reduction of fuel burn and CO_2. Additionally, by an existing FCI, the overall fuel burn can be reduces by 5-10% due to better flight planning, speed management, direct routes, etc. (ACARE, 2010).

6.2 Visions by 2030

Until 2030 the overall vision by using new aeronautical communications technologies in a seamless networking concept is an improved traffic management. The resulting benefits which support the aforementioned visions for 2020 are: less fuel consumption, increase of traffic capacities, less delay in flight operations and better flight planning. Furthermore, instead of stand-alone equipment for each data link, an integrated approach for all communications technologies will reduce weight and power consumption during flights and will benefit in less fuel consumption.

A further goal is the combination of communications and navigation. The new communications systems might be further developed to include a navigation component. Thus, future communications systems could implement alternate positioning navigation and timing (APNT) and act as fallback solutions in the case of a GNSS failure. This will also facilitate smoother transition phases for new system generations due to a better usage of frequency capacities.

6.3 Visions by 2050 and beyond

During the Aerodays 2011 in Madrid, Spain the European Commission released Europe's new vision for aviation by 2050 (Flightpath 2050, 2011). This vision was created by a European High Level Group on Aviation and Aeronautics Research including all key stakeholders of European aviation. The Flightpath 2050 addresses several goals in respect of future communications strategies, for example:

- Travellers can use continuous, secure and robust high-speed communications for added-value applications.
- The transport system is capable of automatically and dynamically reconfiguring the journey within the network to meet the needs of the traveller if disruption occurs.
- An air traffic management system is in place that provides a range of services to handle at least 25 million flights a year of all types of vehicles, (fixed-wing, rotorcraft) and

systems (manned, unmanned, autonomous) that are integrated into and interoperable with the overall air transport system with 24-hour efficient operation of airports.

Besides the Flightpath 2050 there exist also visions of a one pilot cockpit respectively unmanned cockpit which is only feasible with the FCI fully implemented. The necessary ground assistance for a single pilot aircraft or an unmanned aircraft requires highly reliable data communications and high capacity data links which need to be implemented in the final FCI stage.

Furthermore, synergies between sky and sea could be envisioned. This would require a development of a holistic communications infrastructure between aviation and ocean freight / shipping. Since shipping and aviation are using very often the same routes or encounter communications problems in remote areas, this vision envisages a flexible interoperable network between aircraft and ships to enable communication everywhere. Therefore, aviation could support the efficiency of world's largest cargo segment, could also support the reduction of fuel usage (communication of better route planning information), and could support and get communication possibilities in remote areas.

6.4 Readiness level of communications technologies

First studies on seamless aeronautical networking were already done and a proof-of-concept was given, e.g., EU Research Project NEWSKY. A first prototype of such a concept is developed within the EU Research Project SANDRA (SANDRA, 2009). Additionally, an underlying technology of the seamless network is the concept of an aeronautical mobile ad hoc network (MANET). The aeronautical MANET is envisioned to be a large scale multi-hop wireless mesh network of commercial passenger aircrafts connected via long range highly directional air-to-air radio links (cf. Fig. 12)
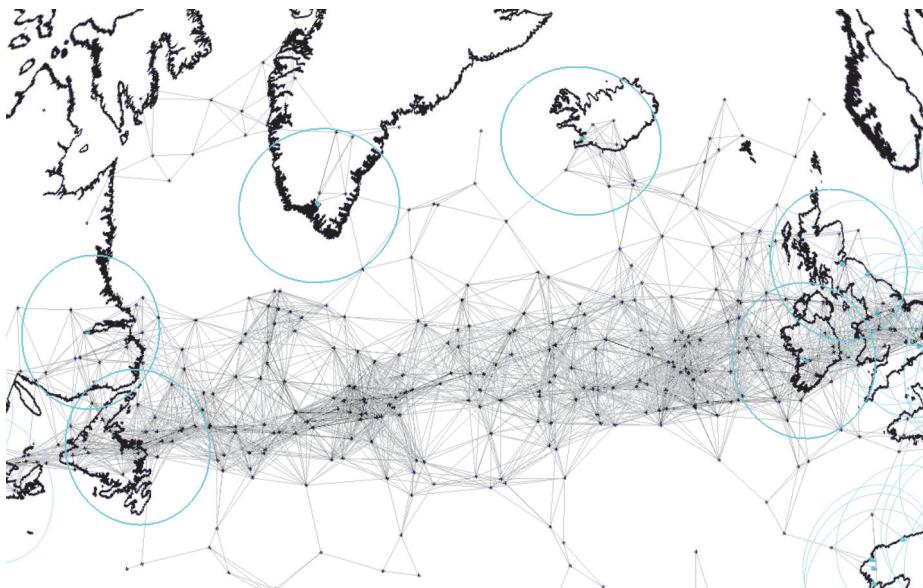
Fig. 12. Example of aeronautical MANET (Medina et al., 2010).

The underlying seamless networking concept is only ready to fully operate by deployment of new data links with higher data rates and flexibilities. An already existing digital data link is VHF Digital Link Mode 2 (VDL2). A high data airport wide data link, namely AeroMACS (EUROCAE WG-82, 2009), is under investigation and also the L-Band Digital Aeronautical Communications System (L-DACS) (Action Plan 17, 2007). Iris, element 10 of the ESA ARTES programme, aims to develop a new air/ground satellite-based solution for the SESAR programme by providing digital data links to cockpit crews in continental and oceanic airspace (Iris, 2009). In addition to the air/ground capability, some of the mentioned data links or unknown future data link technologies could also support air-to-air (A2A), resp. point to point and/or broadcast communications. In the following Table 1 the TRL of these future communications technologies are listed depending on the envisioned decades.

All the aforementioned visions of a fully interconnected world through virtual technologies in 2050 are only feasible by the development and deployment of a FCI based on seamless networking with all communications technologies.

Technology	TRL today	TRL in 2030	TRL in 2050
Seamless Aeronautical Network	3-6	9	9
Aeronautical MANET	2	6	9
VDL2	9	9	9
AeroMACS	5	9	9
L-DACS	4	9	9
Iris	3-4	9	9
A2A	2	6	9
Holistic Network (aviation/shipping)	1	2	6

Table 1. Readiness Level of future aeronautical communications technologies.

7. Conclusions

The International Forum for Aviation Research (IFAR) is a new initiative to connect and represent leading worldwide aerospace research organisations and to allow communication on all global research topics. Climate change is currently the most relevant topic and was the motivation to set up IFAR. However, IFAR also addresses further areas relevant for a future global air transport system (e.g. noise, security, safety, efficient operations). The idea of IFAR was born at the Berlin Summit 2008 where key leaders of 12 international aeronautical research organisations met to address the question of the Air Transport of the Future in the context of climate change. At the second Berlin Summit in 2010 16 international aeronautical research organisations met and eventually set up IFAR. IFAR aims to develop an International Aviation Framework specifically addressing the most important questions for a future global air transport system. In a first stage the Frame work will concentrate on topics related to climate change. Within the next years this Framework is going to be extended by taking the other relevant challenges like noise, safety, security and efficient operations into account. This paper deals with the objectives, state-of-the art and future planning of IFAR. It highlights also first ideas for improved technologies in the area Future Aeronautical Communications for the future. The results of the working groups, the discussions among the participants and the specific actions within the framework development will be regularly updated the IFAR website www.ifar.aero.

8. References

ACARE (2010). Aeronautics and air transport: Beyond Vision 2020 (towards 2050), (June 2010), Available from www.acare4europe.com

Action Plan 17 (2007). Final Conclusions and Recommendations Report, Version 1.1, EUROCONTROL/FAA/NASA, (November 2007)

Berghof, R., Schmitt, A., Eyers, C., et al., 2005. CONSAVE 2050. Final Report. G4MACT-2002-04013 (EU), Cologne.

EREA (2010). EREA – Vision for the Future - Towards the future generation of Air Transport System, (November 2010), Available from www.erea.org

ETP (2010). Energy Technology Perspectives 2010, Scenarios and Strategies to 2050, International Energy Agency (IEA), Available from www.iea.org

EUROCAE WG-82 (2010). WG-82 Mobile Radio Communication Systems: Airport Surface Radio Link (WIMAX Aero), (January 2010), Available from
http://www.eurocae.net/working-groups/wg-list/50-wg-82.html

Flightpath 2050 (2011). Flightpath 2050 - Europe's Vision for Aviation, (March 2011), ISBN 978-92-79-19724-6

IATA (2009). IATA Technology Road Map, 3rd Edition, International Air Transport Association (IATA), (June 2009), Available from www.iata.org

IFAR (2008), International Forum for Aviation Research (IFAR), (May 2008), Available from www.ifar.aero

Iris (2009). Satellite-based communication solution for the Single European Sky Air Traffic Management Research programme - Element 10 of the ESA ARTES programme, Available from www.telecom.esa.int/iris

Medina D., Hoffmann F., Rossetto F., and Rokitansky C.-H. (2010). A Crosslayer Geographic Routing Algorithm for the Airborne Internet, *Proceedings of the IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 2010

NASA (2010). National Aeronautics Research and Development Plan, (February 2010), Available from www.nasa.gov

NextGen (2007). Next Generation Air Transportation System (NextGen), Available from www.faa.gov/nextgen

SANDRA (2009). Seamless aeronautical networking through integration of data links Radios and antennas (SANDRA), Available from www.sandra.aero

SESAR (2009). Single European Sky ATM Research Programme (SESAR) Joint Undertaking, Available from www.sesarju.eu

Szodruch J. and Degenhardt R. (2011a). IFAR- International Forum for Aviation Research, *Aeronautics Days 2011*, Madrid, Spain, 29 March – 01 April, 2011

Szodruch J., Grimme W., F. Blumrich and Schmid R. (2011b). Next generation single-aisle aircraft - Requirements and technological solutions, *Journal of Air Transport Management 17 (2011) 33-39*

The Airborne Internet

Daniel Medina and Felix Hoffmann
German Aerospace Center (DLR)
Oberpfaffenhofen,
Germany

1. Introduction

Mobile communications and internet access are increasingly becoming an essential part of people's lives in today's information society. The growing interest by commercial airlines in providing internet access and cellular connectivity in the passenger cabin has lead to the emergence in recent years of the first satellite-based inflight connectivity providers, including Connexion by Boeing (now defunct), OnAir, AeroMobile, and Panasonic Avionics Corporation. Given the long range of transcontinental air travel, a satellite communications link is the most natural and flexible way to keep the aircraft connected to the ground throughout the flight. Long-distance flights typically traverse oceanic and remote airspace, e.g., large bodies of water, deserts, polar regions, etc., where no communications infrastructure can be deployed on the ground. However, direct air-to-ground (A2G) cellular networks are being deployed (e.g., AirCell in the United States) to provide faster and cheaper access during continental flight.

This Chapter presents the vision of the Airborne Internet, a new paradigm for inflight connectivity based on the concept of *mesh networking* (Akyildiz & Wang, 2005). Airborne mesh networks are self-organizing wireless networks formed by aircraft via direct air-to-air (A2A) radio communication links. Such networks have so far been considered mainly in the context of military aviation (DirecNet, 2007; Bibb Cain et al., 2003).

The concept of the Airborne Internet was first proposed at NASA Langley Research Center's Small Aircraft Transportation System (SATS) Planning Conference in 1999. In one conference session, it was suggested that such a system would require a peer-to-peer communications network among the aircraft. The Airborne Internet Consortium (AIC) formed subsequently to promote and aid in the development of such a system. Consortium members include Aerosat, C3D Aero, and United Airlines.

As shown in Fig. 1, aeronautical mesh networking is envisioned as a means to extend the coverage of A2G access networks offshore to oceanic or remote airspace. By enabling aircraft themselves to act as network routers, an airborne mesh network is formed in the sky, as illustrated in Fig. 2. At any given time, only a fraction of all aircraft are within direct A2G coverage, as they fly over the ground infrastructure deployed on shore. During oceanic flight, the aircraft can stay connected by using the airborne mesh network as a bridge to the ground infrastructure, thus bypassing the costly satellite link. From an airline's perspective, avoiding the satellite link can result in significantly reduced communication costs.

Another potential benefit is reduced latency compared to a geostationary satellite, enabling delay-sensitive applications such as voice and video conferencing. With a geostationary

satellite, there is always a one-way end-to-end propagation delay of approximately 250 ms, required for the signal to travel up and down from the satellite. In the airborne mesh network, lower end-to-end delay guarantees can be provided by making use of appropriate Quality-of-Service (QoS) mechanisms, such as radio resource reservation or packet prioritization.
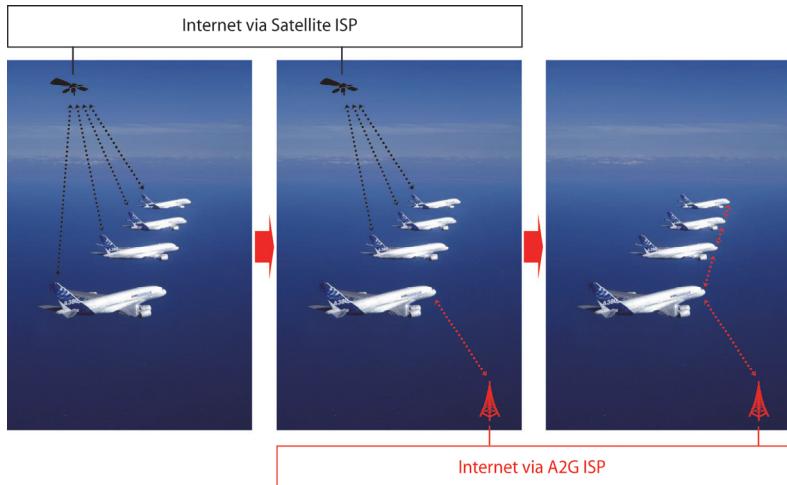

Fig. 1. Evolution from satellite-based to air-to-ground (A2G) inflight connectivity service provision via airborne mesh networking.

Fig. 2. The vision of an Airborne Internet over the North Atlantic.

Initially, an airline may rely only on its own aircraft for mesh networking, since it may be the only airline equipped with the required airborne technology (e.g., antenna, router, etc.). In the longer term, as more and more airlines equip for airborne mesh networking, airline partnerships may be formed to allow their aircraft to mesh together in a single unified cooperative network, building a more richly connected network.

The North Atlantic is the busiest oceanic airspace in the world, and thus constitutes the best candidate scenario for a real deployment of an aeronautical mesh network. In 2007 approximately 425,000 flights crossed the North Atlantic (International Civil Aviation Organization [ICAO], 2008). As a result of passenger demand, time zone differences and airport noise restrictions, much of the North Atlantic air traffic contributes to two major alternating flows: a westbound flow departing Europe in the morning, and an eastbound flow departing North America in the evening. As shown in Fig. 3, the effect of these flows is to concentrate most of the traffic unidirectionally, with peak westbound traffic crossing the 30W longitude between 1130 UTC and 1900 UTC and peak eastbound traffic crossing the 30W longitude between 0100 UTC and 0800 UTC.

Compared to terrestrial mesh networks, the fact that nodes are airborne rather than on the ground makes it possible to communicate over long ranges with unobstructed line-of-sight propagation characteristics. Moreover, nodes are moving at high speeds, giving rise to a rapidly changing network topology.

The maximum communication range in aeronautical mesh networks is constrained by the spherical geometry of the network, as nodes fly very close to the earth surface. The line-of-sight (LOS) communication range is determined by the radio horizon. Within the horizon, atmospheric propagation is essentially subject to free space loss. Attenuation by clouds, rain, etc., can be negligible depending on the frequency spectrum used. Beyond the line-of-sight range, fading due to the earth's obstruction leads to very rapid attenuation (International Telecommunications Union [ITU], 1986).
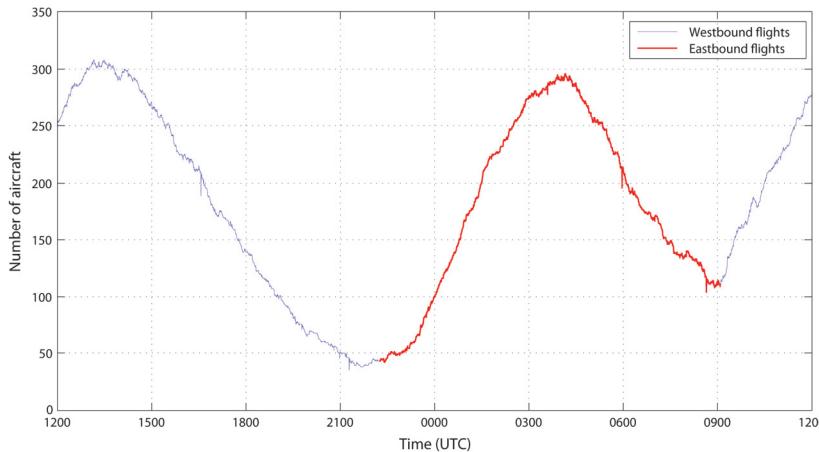

Fig. 3. Number of aircraft in the North Atlantic Corridor throughout the day.

The LOS communication range between two nodes depends on the nodes' flight level and the characteristics of the terrain. In an oceanic environment, the earth surface can be approximated by a perfect sphere, as shown in Fig. 4.
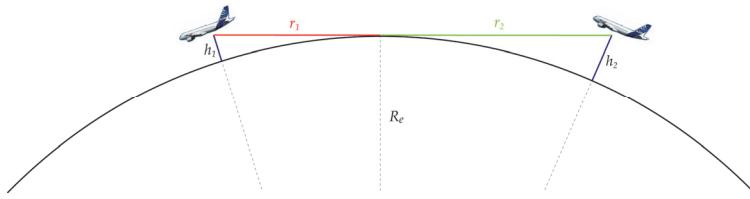
Fig. 4. Line-of-sight (LOS) range.

Given this geometry, Pythagoras' theorem can be used to obtain the LOS communication range

$$r = r_1 + r_2 = \sqrt{h_1^2 + 2R_e h_1} + \sqrt{h_2^2 + 2R_e h_2} \tag{1}$$

where R_e is the earth radius and h_1 and h_2 denote the flight altitude of each aircraft. Typical flight levels for transatlantic flights are between FL310 (31000 ft) and FL400 (40000 ft). For simplicity, we will assume that all aircraft fly at the same altitude h and ground stations are deployed at sea level. Thus, the A2G LOS communication range r_G is given by

$$r_G = \sqrt{h^2 + 2R_e h} \tag{2}$$

and the LOS communication range r between two airborne nodes is

$$r = 2r_G \tag{3}$$

As an example, consider a cruising altitude h = 35000 ft (FL350). Using R_e = 6378.137 km for the earth radius, the LOS communication ranges are $r_G \approx 200$ nmi and $r \approx 400$ nmi. To get an idea of the magnitude of these LOS ranges in relation to airborne node density, Fig. 5 shows a snapshot of transatlantic air traffic at the westbound peak hour (1300 UTC). The air-to-air LOS range, shown by the red circle, covers almost one half of the North Atlantic airspace.
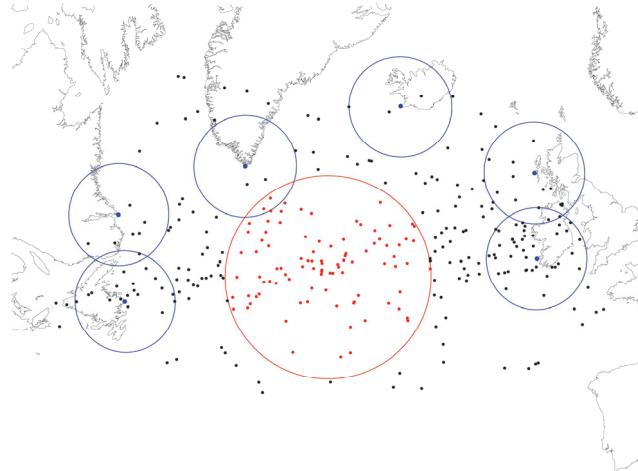

Fig. 5. Air-to-air (red) and air-to-ground (blue) LOS range at FL350.

Of course, the nominal communication range may be smaller than the LOS range, depending on the transmit power, the characteristics of the antenna, the modulation (data rate) used for transmission, and the target bit error rate (BER). The AeroSat Corporation, together with the U.S. Federal Aviation Administration (FAA), has performed flight trials with mechanically steered Ku-band antennas, demonstrating A2G link data rates of up to 45 Mbps over 150 nautical miles with a BER of 5.10^{-5} (McNary, 2007). Looking forward, we believe smart antennas to be the most appropriate technology for broadband airborne mesh networking, since they allow a node to quickly change the direction(s) in which it transmits/receives to/from its various neighbors and optimize the signal-to-interference ratio at the receiver (Bhobe & Perini, 2001).

Broadband airborne mesh networks require a medium access control (MAC) protocol capable of handling high traffic loads in the network and providing QoS guarantees to communicating nodes. Carrier sense multiple access (CSMA) techniques are inappropriate in this environment, given the long propagation delay (a couple milliseconds) and the directional nature of radio transmissions. Aircraft are equipped with GPS for navigation purposes, and this provides a global time reference that can be exploited for synchronization among network nodes, e.g., to schedule collision-free transmissions in a time division multiple access (TDMA) fashion (Nelson & Kleinrock, 1985).

In this Chapter, we propose a novel routing strategy that takes into account the specific nature of aeronautical mesh networks. A number of observations have guided our design. The airborne mesh network is connected to the ground at potentially multiple geographically distributed access points (Internet Gateways) via a rapidly changing number of short-lived bandwidth-limited A2G links, through which all internet traffic enters/leaves the airborne leaf network. We envision passengers consuming (rather than producing) great amounts of information, resulting in a considerable aggregate downstream traffic volume being delivered to the airborne network from the Internet Gateways. Thus, the Internet Gateways pose a capacity bottleneck, limiting the maximum bandwidth that can be offered to the aircraft. At any given time, an aircraft may be able to reach multiple Internet Gateways via a number of disjoint paths. This path diversity can be exploited to reduce congestion at the bottleneck A2G links. Our proposed strategy, *Geographic Load Share Routing* (GLSR), exploits the aircraft's position information (e.g., made available through GPS) together with buffer size information to fully exploit the total A2G capacity available at any time to the airborne network by balancing the aggregate traffic load among all A2G links.

The remainder of this Chapter is structured as follows. Section 2 provides references to related work. Section 3 describes our network model, including the antenna and interference model used in our simulations. In Section 4, we formulate a joint routing and scheduling optimization problem to minimize the average packet delay in the network. Section 5 briefly describes the link scheduling algorithm used to assign capacity to network links. Our proposed routing strategy is presented in Section 6, followed by a maximum throughput analysis in Section 7. Our simulation results are presented and discussed in Section 8. Finally, Section 9 concludes the Chapter.

2. Related work

Although a great number of routing protocols have been proposed for wireless mesh networks (Akyildiz & Wang, 2005), to the best of our knowledge none of them has been designed with the specific goal of aeronautical mesh networking in mind, and therefore they

do not exploit the distinct characteristics of this environment. Only very recently has some attention been drawn to the application of multihop wireless networking to aviation (Sakhaee & Jamalipour, 2006; Sakhaee et al., 2006; Iordanakis et al., 2006; Tu & Shimamoto, 2009). However, these authors have a different focus and relatively simple network models. In previous work (Medina et al., 2008a; Medina et al., 2008b), we conducted simulations of realistic air traffic to study the feasibility and characterize the topology of such networks. For an excellent survey on geographic routing, see (Mauve et al., 2001). (He et al., 2003) proposed SPEED, a stateless protocol for real-time communication in wireless sensor networks. SPEED uses a geographic forwarding strategy similar to our own, which we already presented in (Medina et al., 2010) and forms part of the overall routing strategy presented here. Internet Gateway selection in mobile ad hoc networks is addressed in (Sun et al., 2002; Huang et al., 2003; Brännström et al., 2005; Ahn et al., 2005). Selection strategies generally assume omnidirectional transmissions and IEEE 802.11 as the underlying medium access technology.

3. Network model

As shown in Fig. 6, the network consists of an airborne segment (the airborne mesh network) and a ground segment (the A2G access network). At any time, the airborne network consists of a variable number N of mobile nodes (aircraft), whereas the ground segment is composed of a fixed number M of geographically distributed stationary ground stations (Internet Gateways), assumed to be operated by an A2G communications provider. A particular node in the network is uniquely identified by its number $i \in \{1, ..., N+M\}$.
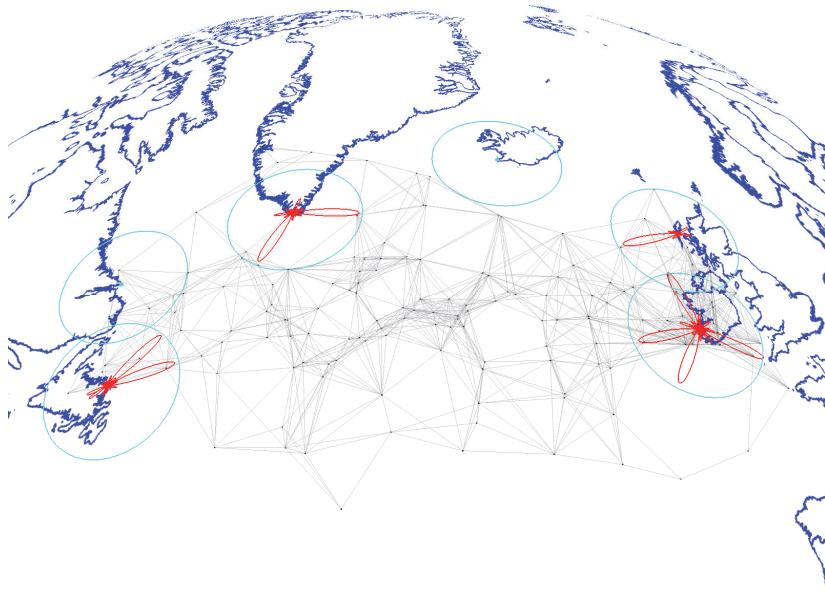

Fig. 6. Network model.

Direct communication from node i to node j is represented by the directed link (i,j), $i \neq j$. A link (i,j) exists if a sufficiently low bit error rate can be achieved in the absence of multiple access interference. In the absence of interference, the bit error rate depends on the signal-to-noise ratio (SNR) at the receiving end of the link. For simplicity, we assume that all nodes use the same transmit power, high enough for a link to be feasible with any other node within the radio horizon, given by (2)-(3).

3.1 TDMA medium access control

All nodes (aircraft and ground stations) are assumed to use half-duplex transceivers on the same carrier frequency (common channel) and are assumed to be synchronized to a common time reference, e.g., by means of GPS. To avoid multiple access interference, link transmissions are scheduled in a TDMA fashion. The time domain is divided into repeating frames of size T time slots, each with a duration T_s long enough to transmit one packet. Transmissions start and end within a slot. The TDMA schedule specifies a link's activation pattern over the frame, that is, during which time slots it can transmit a data packet. The size of a packet corresponds to the duration of a time slot minus the appropriate guard time, required to offset the varying geographic distances between nodes.
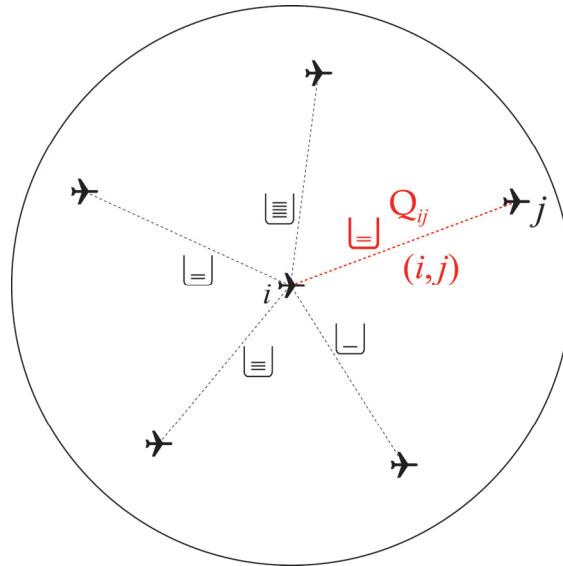
Fig. 7. A node's transmission queues.

We denote by N_i the set of all one hop neighbors of node i. As shown in Fig. 7, every node i has an outgoing link (i,j) with each neighbor $j \in N_i$, with an associated transmission queue Q_{ij} where arriving packets are buffered while they wait for transmission over link (i,j). For each queue Q_{ij} in the network, the packet arrival rate λ_{ij} is computed at the beginning of each frame n using an exponentially weighted moving average, given by

$$\lambda_{ij}^{(n)} = (1-\kappa)\lambda_{ij}^{(n-1)} + \kappa\Lambda_{ij}^{(n-1)} \qquad (4)$$

where $\Lambda_{ij}^{(n)}$ is the number of packet arrivals at Q_{ij} during frame n. The moving average is used to smooth out short-term fluctuations in the arrival rate. The arrival rate λ_{ij} of each link (i,j) is used by the traffic sensitive link scheduling algorithm (described in Section 5) to assign time slots to links proportionally to their traffic demand.

Let h_{ij} denote the number of time slots currently assigned to link (i,j). The capacity of link (i,j) is given by

$$c_{ij} = h_{ij} / T \tag{5}$$

where T is the frame length (number of slots). Thus, the capacity of a link is given by the fraction of time slots in the frame that it has been assigned for transmission by the link scheduling algorithm. Note that, in general, $c_{ij} \neq c_{ji}$.

3.2 Antenna and interference model

As shown in Fig. 8, we use a uniform circular array antenna model, whereby only the signal phases (not the amplitudes) of the array elements are controlled to steer the main beam toward the strongest signal path, i.e., the line of sight. Beam steering is used in both transmission and reception. In addition, we assume that the uniform circular array can form up to K independent beams *simultaneously* in arbitrary directions for concurrent packet transmission/reception and can quickly reconfigure the directions in which it transmits or from which it receives at the beginning of every time slot (fast beam switching). The antenna pattern of a uniform circular array can be found in (Balanis, 2005; Moser, 2004). The half-power beamwidth ψ and the main lobe antenna gain depend on the number of array elements n_{elem}.

| n_{elem}	Gain (dB)	Pattern
4	6.0	
8	9.0	
16	12.0	
32	15.0	
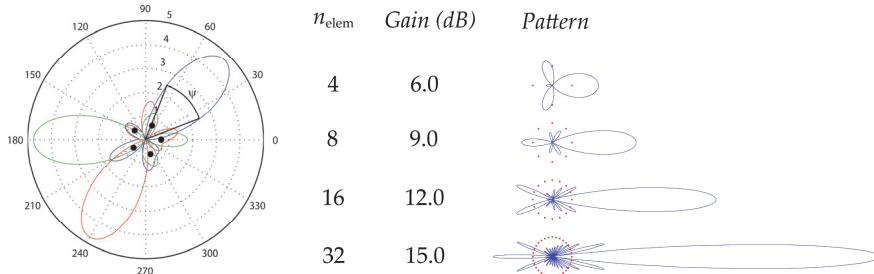
Fig. 8. Multibeam uniform circular array antenna azimuthal radiation patterns.

We define the maximum interference distance ρ as the distance from the transmitter beyond which interference is assumed to be zero. As with the LOS communication range, the maximum interference distance will depend on the altitudes of the transmitter and the receiver. Thus, we distinguish between the maximum A2G interference distance ρ_G and the maximum A2A interference distance ρ. We use the values $\rho_G = 225$ nmi and $\rho = 450$ nmi.

For each communication link (i,j) (see Fig. 9), the signal-to-interference ratio (SIR) in a given slot s is computed as

$$\Gamma_{ij}[s] = \frac{G_{ij}(\theta_{ij})G_{ji}(\theta_{ji})d_{ij}^{-\alpha}}{\sum_{\forall(k,l)\neq(i,j)} G_{kl}(\theta_{kj})G_{ji}(\theta_{jk})d_{kj}^{-\alpha}z_{kl}[s]} \tag{6}$$

where G_{ij} is the antenna radiation pattern used by node i to transmit to node j, θ_{ij} is the azimuthal angle to node j as seen from node i, d_{ij} is the line-of-sight distance between nodes i and j, α is the path loss exponent (we assume $\alpha = 2$), and

$$z_{kl}[s] = \begin{cases} 1, & \text{if link } (k,l) \text{ is scheduled in slot s and } d_{kj} < \rho_{(G)} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$
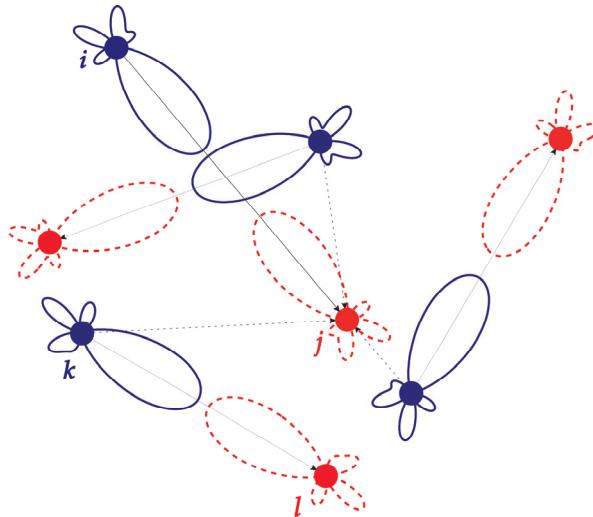

Fig. 9. Example network illustrating our SIR-based interference model (solid lines represent communication links, dotted lines represent interference links).

Simultaneous link activation in a given time slot is limited by the following constraints:

(c_1) Half duplex operation: A node cannot transmit and receive simultaneously.

(c_2) A node may activate at most K outgoing (transmit mode) or incoming (receive mode) links simultaneously.

(c_3) The signal-to-interference ratio at all scheduled receivers must be above a specified communication threshold γ_o.

We assume that link (i,j) can transmit a packet without error in slot s if $\Gamma_{ij}[s] > \gamma_o$.

4. Joint routing and scheduling optimization

In order to determine the maximum network performance in terms of throughput or delay that can be achieved in the aeronautical networks considered here, it is useful to formulate an optimization problem minimizing the average packet delay in the network, subject to constraints that require the existence of a feasible schedule. Intuitively, minimizing the packet delay is a reasonable design goal, since this metric is directly related to the quality of service that is perceived by a user. We denote the set of all traffic flows in the network as F. A flow (p,q) in F is defined by its source and destination nodes and is associated with a target data rate R_{pq}, given in slots per frame. We introduce the variables

$$u_{ij}[\mathrm{s}] = \begin{cases} 1, & \text{if link } (i,j) \text{ is scheduled in slot s} \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

and

$$l_{ij,pq} = \begin{cases} 1, & \text{if link } (i,j) \text{ carries traffic for flow } (p,q) \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

The average packet delay on a wireless link can be approximated by the time that the packet must wait until a transmission opportunity for this link, i.e., until a time slot allocated to this link arrives in the schedule (time slot offset), plus the transmission time itself. Assuming that the slots for a link are distributed at uniform intervals in the TDMA frame, the average delay on link (*i,j*) can be expressed as

$$D_{ij} = T_s \left(1 + \frac{T}{2 \sum_{\mathrm{s}=1}^{T} u_{ij}[\mathrm{s}]} \right), \tag{10}$$

and the average packet delay in the network is given by

$$\widehat{D} = \frac{1}{\sum_{(p,q) \in F} R_{pq}} \left(\sum_{(p,q) \in F} R_{pq} \sum_{(i,j)} l_{ij,pq} D_{ij} \right). \tag{11}$$

Note that the average packet delay depends on both the routing variables $l_{ij,pq}$ and the scheduling variables $u_{ij}[\mathrm{s}]$. When the traffic demand is known, the link delay is a convex function of the scheduling variables. Unfortunately, the joint routing and scheduling delay minimization problem is non-convex, so that the global optimum cannot be found in general. Therefore, we split the problem up into two steps: First, a minimization of the weighted hop count (mWHC), subject to constraints requiring a feasible schedule; second, minimization of the average flow delay (mAFD), given the link loads resulting from the solution of the first step. The problem formulation is summarized in the table below.

The coefficients w_{ij} in the objective function (12a) allow links to be assigned different weights. For example, a higher weight can be given to satellite links than to A2A links in order to avoid the high delay and cost that are typically associated with satellites. The first two routing constraints (12b), (12c) ensure that flows begin at their source and end at their destination, respectively. The third constraint (12d) ensures that traffic flow is conserved at intermediate nodes. The last three constraints concern the scheduling. The first (12e) ensures half duplex operations, the second (12f) enforces that the capacity of each link is sufficient to carry the link's traffic load, and the final constraint (12g) ensures that the SIR of all active links is above the SIR threshold required for error free communication. The mAFD problem does not require the routing constraints, since the routing has already been decided in Step 1. The constraints are the same as the scheduling constraints for mWHC.

Unfortunately, the applicability of this approach is limited to very small networks due to the large number of integer variables. A more efficient, but suboptimal, approach based on

Step 1:   Step 2:
Weighted Hop Count Minimization   Average Flow Delay Minimization

$$\min \sum_{(i,j)} w_{ij} \sum_{(p,q) \in F} R_{pq} l_{ij,pq} \qquad (12a) \qquad\qquad \min \sum_{(p,q) \in F} \sum_{(i,j)} l_{ij,pq} D_{ij} \qquad (13a)$$

s.t.
$$\sum_{j} l_{pj,pq} = R_{pq} \quad \forall (p,q) \qquad (12b) \qquad \text{s.t.} \qquad \sum_{i} u_{ij}[s] + \sum_{i} u_{ji}[s] \le 1 \quad \forall j,s \qquad (13b)$$

$$\sum_{i} l_{iq,pq} = R_{pq} \quad \forall (p,q) \qquad (12c) \qquad\qquad \sum_{s} u_{ij}[s] \ge \sum_{(p,q) \in F} l_{ij,pq} R_{pq} \quad \forall (i,j) \qquad (13c)$$

$$\sum_{i} l_{ik,pq} = \sum_{i} l_{ki,pq} \quad \forall k \qquad (12d) \qquad\qquad \Gamma_{ij}[s] \ge \gamma_0 u_{ij}[s] \quad \forall (i,j),s \qquad (13d)$$

$$\sum_{i} u_{ij}[s] + \sum_{i} u_{ji}[s] \le 1 \quad \forall j,s \qquad (12e)$$

$$\sum_{s} u_{ij}[s] \ge \sum_{(p,q) \in F} l_{ij,pq} R_{pq} \quad \forall (i,j) \qquad (12f)$$

$$\Gamma_{ij}[s] \ge \gamma_0 u_{ij}[s] \quad \forall (i,j),s \qquad (12g)$$

Genetic Algorithms (GAs) has been described in (Hoffmann et al., 2011). In contrast to the mathematical programming approach, the GA does not need to be solved anew with each change in the topology, but can be run "on the fly", while the network is moving. In addition, the GA can also be successfully applied to non-convex problems, allowing a direct minimization of the average packet delay. In the proposed GA, a random path to a random gateway is selected for each flow when an individual of the initial population is created. The subsequent operations of recombination and mutation may modify the scheduling of links, the routing of a single flow, or exchange entire paths between individuals of the population. In small networks, the GA provides performance results similar to what can be achieved by means of the mWHC/mAFD approach. It has been shown in (Hoffmann et al., 2011) that the GA easily outperforms hop count based routing and gateway selection in larger networks. However, both of these approaches are mainly of interest to determine performance bounds of the network. They require global knowledge of the traffic demands and aircraft positions at a centralized processor. For practical purposes, it is evident that distributed routing and scheduling algorithms are required. These will be addressed in the following.

5. Distributed STDMA link scheduling

In order to assign capacity to links proportionally to their traffic load, we use the traffic sensitive STDMA link scheduling algorithm proposed in (Grönkvist, 2005). In this section, we provide a brief summary of the essential aspects of the algorithm. For a detailed description, see (Grönkvist, 2005).
The priority of a link (i,j) is defined as

$$p_{ij} = \lambda_{ij} / h_{ij} \qquad (14)$$

where λ_{ij} is the packet arrival rate at Q_{ij} (in packets/frame) given in (4) and h_{ij} is the number of slots currently assigned to link (i,j). The link priorities are used by the link scheduling algorithm to provide fairness among links competing for radio resources in the network.

The local neighborhood L_{ij} of link (i,j) is defined as the set of all other links (k,l) in the network whose transmitter k is within interference distance of j and/or whose receiver l is within interference distance of i, i.e.,

$$\mathcal{L}_{ij} = \left\{ (k,l) \ : \ d_{kj} < \rho \right\} \ \cup \ \left\{ (k,l) \ : \ d_{il} < \rho \right\}$$ (15)

The distributed STDMA algorithm consists of the following steps:
1. Nodes that have entered the network exchange local information with their neighbors.
2. The link with highest priority in its local neighborhood assigns itself a time slot.
3. The local schedule is then updated within the local neighborhood of the link, and a new link has highest priority.

This process (2.-3.) is continued until all slots are occupied, i.e., there are no available slots to assign. In this way, link priority decides in which order links may attempt to assign themselves a time slot. If no slot is available, the link may steal an allocation from a lower priority link in the local neighborhood.

A time slot assignment is maintained for as long as possible until either it can no longer be used reliably or it is stolen by a higher priority link. Node movement will cause topological changes and modify the interference geometry, so that allocations that were compatible at one time cease to be so at a later time. Every node continuously monitors the SIR of its incoming links and drops any allocations whose SIR has become lower than the communication threshold γ_o, notifying its local neighborhood about the deallocation.

6. Geographic load share routing

In the Airborne Internet, every ground station on shore acts as an Internet Gateway (IGW). IGWs periodically announce their existence and geographic location via IGW advertisements. An aircraft may receive advertisements from potentially multiple IGWs, but at any time uses only one of them as its default IGW for all A2G communications, which is kept up-to-date on the aircraft's current position. An aircraft only forwards to its neighbor aircraft advertisements originating from its default IGW. Whenever appropriate, a handover procedure is used to change an aircraft's default IGW.

Consider, as shown in Fig. 6, a snapshot of the network topology at a given time. We make the following assumptions in the sequel:
•   Only downstream traffic is considered. In general, passengers are much more likely to consume than to produce information, so the bulk of the data will flow from the Internet to the airborne network.
•   Every aircraft has the same data traffic demand λ.
•   The airborne network is not partitioned, i.e., there exists at least one path between any two aircraft.

Let L_G denote the set of all A2G links (i,j) from a ground station i to an aircraft j.[1] The maximum instantaneous per-aircraft throughput theoretically achievable is then given by

$$\mu_{max} = \frac{1}{N} \sum_{(i,j) \in L_G} c_{ij} = \frac{C}{N}$$ (16)

---

[1] We use the acronym A2G, rather than G2A, although we are referring to the directed links from the ground to the aircraft.

where N is the number of aircraft forming the airborne network, C denotes the total A2G capacity available to the airborne mesh network, and c_{ij} is given by (5).

In order to fully exploit the total A2G capacity available at any given time, we propose a novel routing scheme, *Geographic Load Share Routing* (GLSR), to balance the traffic load among all A2G links. GLSR consists of two separate strategies:

- a *forwarding* strategy, enabling every intermediate node to choose the next hop on a packet-by-packet basis using only position and buffer size information local to the forwarding node, and
- a *handover* strategy, enabling the access network to control which aircraft is associated with which IGW at any time, based on geographic proximity and a measure of IGW congestion.

6.1 GLSR forwarding strategy

The GLSR forwarding strategy works as follows. Consider a packet arriving at node i with destination m, as shown in Fig. 10.
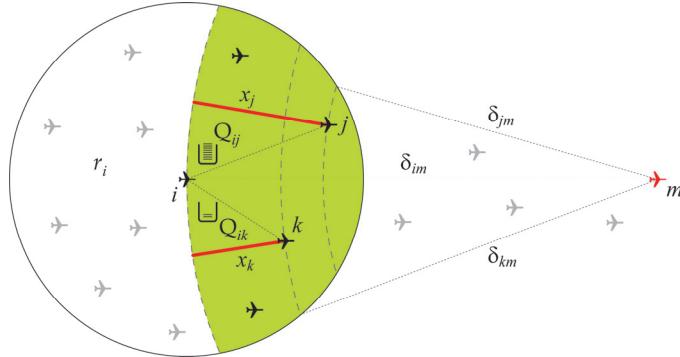

Fig. 10. Airborne forwarding.

The packet's *advance* toward m if forwarded to neighbor k, denoted by x_k, is defined as

$$x_k = \delta_{im} - \delta_{km} \tag{17}$$

where δ_{ij} denotes the (great circle) distance between nodes i and j. The standard geographic forwarding strategy, known as *greedy forwarding* (see, e.g., (Mauve et al., 2001)), chooses as the next hop for a packet the neighbor that is geographically closest to the packet's final destination. Thus, greedy forwarding places a packet arriving at node i with destination m in transmission queue Q_{ij} such that

$$. x_j = \max_{k \in \mathcal{N}_i}\{x_k\}, \ x_k > 0 . \tag{18}$$

If the packet arrival rate at Q_{ij} is higher than the capacity assigned to link (i,j), i.e., $\lambda_{ij} > h_{ij}$, the buffer size will grow, since packets arrive at a greater rate than they can be transmitted. This will lead to increased queueing delay of packets, and may eventually result in packets being dropped due to buffer overflow, unless link (i, j) is able to obtain additional slots. We define a packet's *speed of advance* toward destination m for neighbor k as

$$v_k = \frac{x_k}{n_k + 1} \qquad (19)$$

where n_k is the number of packets in Q_{ik} upon arrival. GLSR places a packet arriving at node i with destination m in Q_{ij} such that

$$v_j = \max_{k \in \mathcal{N}_i} \{v_k\}, \; v_k > 0 \qquad (20)$$

If the destination m is a neighbor, the packet is simply placed in Q_{im}. Thus, GLSR chooses the neighbor which maximizes the ratio given by the packet's advance, as used in greedy forwarding, over the packet's queueing delay, as in a Join the Shortest Queue (JSQ) discipline. As information about the buffer size is local to the forwarding node, it does not need to be sent over the channel, thus introducing no additional overhead.

Note that, in order to guarantee loop free routing, only neighbors with positive speed of advance are considered for load sharing (shaded area in Fig. 10). However, there is a chance that no neighbor aircraft is geographically closer to the packet's destination than the forwarding node. This situation is known in the literature as a communications void (for a survey of void handling techniques, see (Chen & Varshney, 2007)). We note, however, that given the airborne node density in the region of interest for the Airborne Internet, the line-of-sight radio horizon between two airborne nodes (in the order of 400 nautical miles at 35000 ft) and the spatiotemporal nature of transatlantic air traffic patterns, such a communications void in any direction of interest is extremely unlikely.

In the special case where the forwarding node i is the Internet Gateway itself (where the downstream packet originates), as shown in Fig. 11, the packet's speed of advance toward destination m for neighbor k is given by

$$v_k^G = \frac{x_k + r_G}{n_k + 1} \qquad (21)$$

where r_G is the A2G communications range. In this way, all aircraft within the IGW's radio horizon, including those further from the destination than the IGW itself, yield a positive speed of advance. Once a packet enters the airborne network, it may not be forwarded back to the ground, thus precluding a routing loop.
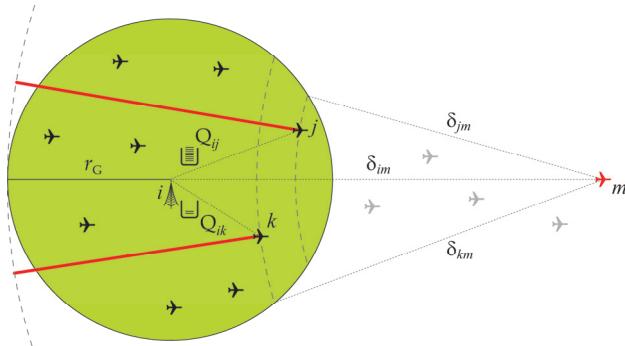

Fig. 11. Ground station forwarding.

6.2 GLSR handover strategy

In order to increase per-aircraft bandwidth, an inflight connectivity provider will likely deploy an A2G access network composed of geographically distributed ground stations along the coast, at appropriate locations dictated by the expected transoceanic air traffic patterns of its customer airlines. The total data traffic demand in the airborne mesh network can then be better accommodated by sharing the load among multiple IGWs.

A trivial approach to the Internet Gateway assignment problem is shown in Fig. 12. Nodes are assigned to the geographically closest (topologically reachable) IGW. The dotted lines represent the Voronoi diagram corresponding to the set of points where the IGWs are located. Each Voronoi cell V_i represents the area formed by all points on the sphere whose geographically closest IGW is i. All aircraft within V_i are served by IGW i. Whenever an aircraft crosses a cell boundary, say, from V_i to V_j, a handover procedure is performed between the aircraft and the access network to transfer all A2G communications for that aircraft from IGW i to IGW j.
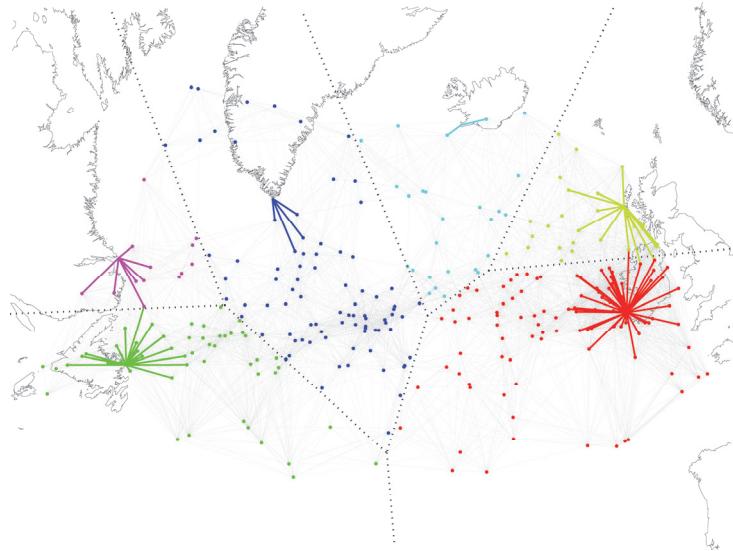

Fig. 12. Internet Gateway assignment based on geographic proximity (Voronoi diagram).

The proximity criterion ignores two important aspects:

- The spatiotemporal distribution of traffic demand in the airborne mesh network. At any given time, the aggregate traffic demand from all airborne nodes in a Voronoi cell may vary greatly among different cells, e.g., the number of nodes V_k flying within each Voronoi cell V_k can be very different.

- The total A2G capacity $C_k = \sum_{l \in \mathcal{N}_k} c_{kl}$ at each IGW k. A richly connected IGW may be able to serve a larger number of users, e.g., by performing load sharing among A2G links. Compare the IGWs in Ireland (over forty A2G links) and Iceland (just two A2G links) in Fig. 12.

A simple way to address these two important aspects together is to consider the impact of an imbalance between A2G demand and A2G capacity on an IGW's transmission buffers.

Consider IGW k and let $\widehat{Q}_{kl}$ denote the average buffer size of transmission buffer Q_{kl}, i.e., the average number of packets waiting for transmission over A2G link (k,l). By virtue of the GLSR forwarding strategy described in the previous section, A2G traffic load will be shared among all A2G links at IGW k. In order to characterize quantitatively the ratio of A2G demand to A2G capacity, we define the *congestion* at IGW k as the maximum average buffer size among all its A2G links, i.e.,

$$\Omega_k = \max_{l \in \mathcal{N}_k} \left\{ \widehat{Q}_{kl} \right\} \tag{22}$$

The objective is to balance traffic load among IGWs in order to prevent unnecessary congestion at an IGW while other IGWs have free available capacity. This requires a handover management strategy that takes into account not only the geographic coordinates of the airborne nodes, but also the congestion measure at each IGW, as defined in (22). To achieve this, GLSR relies on a centralized Internet Gateway handover manager in the access network, which is assumed to know the current geographic coordinates (θ_m, φ_m) of every airborne node m in the network, as well as the congestion measure Ω_k for each IGW k. For every airborne node m, we define its *congestion distance* to Internet Gateway k as

$$\Delta_{km} = \delta_{km} \left(1 + \Omega_k \right) \tag{23}$$

The GLSR handover strategy works as follows. Every τ_h seconds (*handover period*), the IGW handover manager computes for every aircraft m (currently associated with IGW i)

- its current congestion distance Δ_{im}
- the IGW j at minimum congestion distance, i.e., satisfying

$$\Delta_{jm} = \min_k \left\{ \Delta_{km} \right\} \tag{24}$$

Note that, by virtue of (24), we have $\Delta_{im} \geq \Delta_{jm}$. If $i = j \;\forall m$, no handover is required.

Otherwise, the aircraft h with greatest metric ratio, i.e., satisfying

$$\frac{\Delta_{ih}}{\Delta_{jh}} = \max_m \left\{ \frac{\Delta_{im}}{\Delta_{jm}} \right\} \tag{25}$$

performs a handover from IGW i to IGW j.

Thus, GLSR periodically checks whether any airborne node can enjoy a shorter congestion distance to the access network, given the current geographic distribution of the airborne network and the current congestion situation at the access network. If every aircraft is associated with the IGW at minimum congestion distance, no handover is required. Otherwise, the aircraft which can benefit most from a handover (i.e., has the greatest metric ratio, as given in (25)) performs a handover to the IGW at minimum congestion distance.

7. Maximum throughput analysis

Consider the following three routing schemes:

[G+V] *Greedy forwarding with fixed Voronoi cells*. No load sharing takes place. Packets are always forwarded to the next hop that is closest to the final destination. An aircraft chooses as its default IGW the geographically closest one.

[S+V] *Speed of advance forwarding with fixed Voronoi cells.* The speed of advance metric is used to balance load among A2G links at each IGW, but no load sharing is performed among IGWs, i.e., each aircraft is associated with the geographically closest IGW.

[S+H] *Speed of advance forwarding with cell breathing.* Load sharing takes place among A2G links, via the speed of advance metric, and among IGWs, via the congestion distance metric. The maximum per-node throughput with greedy forwarding is given by

$$\mu_{G+V} = \min_{(i,j)\in\mathcal{L}_G}\left\{\frac{c_{ij}}{G_{ij}}\right\} \tag{26}$$

where G_{ij} denotes the number of airborne nodes in Voronoi cell V_i whose traffic is routed via A2G link (i,j).

On the other hand, when packets are forwarded according to their speed of advance, all A2G links available at IGW k may be used to route packets to any of the V_k destination aircraft within Voronoi cell V_k. Which specific A2G link is used to transmit a packet will depend on the position of the destination aircraft and the state of the multi-queue system at the IGW upon arrival. Thus, the total A2G capacity $C_k = \sum_{l\in N_k} c_{kl}$ is shared equally by all V_k aircraft in cell V_k. The maximum per-node throughput is therefore given by

$$\mu_{S+V} = \min_{k}\left\{\frac{C_k}{V_k}\right\} \tag{27}$$

The GLSR handover strategy effectively adapts the size of each cell based on the congestion measure at each IGW, giving rise to cell breathing. A cell experiencing congestion will become increasingly unattractive to nodes close to the cell boundary, causing them to perform handovers to neighboring cells with lower congestion. Thus, the cell in question will effectively shrink. As traffic demand increases, the combined effect of both geographic load sharing strategies is such that cells with higher total A2G capacity will swallow nodes from congested cells with lower A2G capacity, until a congestion equilibrium is found among neighboring cells. In saturation, the number of nodes in cell k, denoted by N_k, will be roughly proportional to the total A2G capacity C_k available at IGW k. Thus, the ratio C_k/N_k will be approximately the same for every cell k, and the maximum per-aircraft throughput will approach the theoretical maximum given in (16), as

$$\mu_{S+H} = \min_{k}\left\{\frac{C_k}{N_k}\right\} \approx \frac{C}{N} = \mu_{max} \tag{28}$$

Thus, through the combination of both strategies we fully exploit the total A2G capacity C available at any given time to the airborne network via all A2G links.

8. Simulation results

In order to assess the performance of our routing strategy in a realistic aeronautical scenario, we have implemented our network model in the OMNeT++ simulation framework (OMNeT++, 2011). The simulated scenario consists of six Internet Gateways, placed as shown in Fig. 6. We generate air traffic according to the airline flight schedule database

published by the International Air Transport Association (IATA), containing the departure and destination airports and schedules of all commercial airlines worldwide in operation today (IATA, 2007). We simulate a 24 hour time window (starting at 1200 UTC) corresponding to an average day (in terms of air traffic volume). Flight trajectories are approximated by great circle arcs between departure and destination airports. We assume a 50% equipage level and thus generate each transatlantic flight with a probability of 0.5.

All aircraft are assumed to fly at the same altitude of 35000 ft, resulting in an A2G range r_G = 200 nmi. The airborne topology is controlled by every aircraft by applying the distributed Cone-Based Topology Control (CBTC) algorithm proposed in (Li et al., 2005). For any given aircraft i, the set of neighbors N_i is formed by all nodes within the minimum range r_i, with $r_G \le r_i \le 2r_G$, such that every cone of 120° contains at least one neighbor aircraft.

Internet traffic is generated at each IGW k based on a Poisson traffic model with mean value $N_k \lambda$ packets/sec, where N_k is the number of aircraft served by IGW k and λ is the per-aircraft traffic demand, which is the same for all aircraft. Each new packet's destination is chosen randomly among all aircraft in the IGW's aircraft set.

Our simulation settings are summarized in Table 1.

r_G	200 nmi
r_i	$r_G \le r_i \le 2r_G$
ρ_G	225 nmi
ρ	450 nmi
T	25 slots
T_s	10 ms
K	8 beams
n_{elem}	32
τ_h	5 s

Table 1. Simulation settings.

8.1 Results with idealized wireless channel access

In order to more clearly demonstrate the load sharing behavior of GLSR, we first abstract away the complexity of the underlying wireless channel and assume that every link can transmit simultaneously without interference or half-duplex constraints. The scheduling algorithm described in Section 5 is turned off and every link is allowed to transmit in every time slot, resulting in a uniform link capacity c_{ij} = 1 packets/slot for every link (i,j).

8.1.1 Maximum instantaneous throughput

Fig. 13 shows the maximum per-aircraft throughput over a period of 24 hours for the three routing schemes defined in Section 7. To obtain the maximum instantaneous per-node throughput, denoted by Θ, the per-aircraft traffic demand λ is incremented (decremented) at the beginning of each time frame n according to

$$\lambda_n = \lambda_{n-1} + \Delta\lambda \left(1 - 2\frac{\max_k \Omega_k}{Q_{max}} \right) \tag{29}$$

with the values $\Delta\lambda$ = 0.1 packets/sec, maximum buffer size Q_{max} = 20 packets and Ω_k as defined in (22). Packets arriving at a full buffer are dropped.

The rationale for (29) is that the Internet Gateway with maximum congestion level $\max_k \Omega_k$ represents the traffic bottleneck. Whenever $\max_k \Omega_k$ < $Q_{max}/2$, the per-aircraft traffic demand λ is uniformly increased for all airborne nodes. Whenever $\max_k \Omega_k$ > $Q_{max}/2$, λ is decreased. As a result, the traffic demand stabilizes at any given time around a value such that $\max_k \Omega_k$ $\approx Q_{max}/2$, which is used as the maximum throughput criterion. The throughput curves Θ_{G+V}, Θ_{S+V} and Θ_{S+H} give the real throughput obtained by dividing the number of successfully delivered packets by the number of aircraft, with one data point generated every 10 seconds.
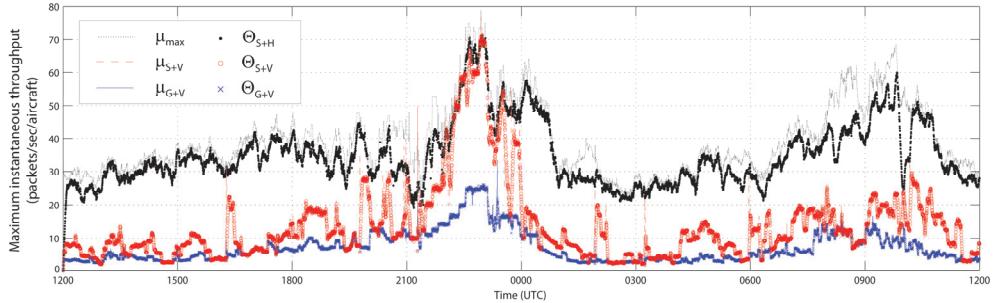

Fig. 13. Maximum instantaneous per-aircraft throughput.

The G+V routing scheme, akin to a shortest path routing strategy, does not exploit the A2G path diversity present in the network, and leads to congestion at low demand levels, since a single A2G link is responsible for carrying traffic to many aircraft, while most other A2G links are underutilized. On the other hand, speed of advance forwarding balances traffic load among all of an Internet Gateway's A2G links, exploiting its full capacity. But if the Internet Gateway has only a few A2G links (in the worst case, a single link) and is geographically closest to a big portion of the airborne network, there is little gain to be expected from the GLSR forwarding strategy alone (S+V routing scheme). As an example, consider the Greenland IGW at 1300 UTC (see Fig. 15).

The S+H routing scheme yields a throughput Θ_{S+H} very close to the theoretical maximum μ_{max}, except at certain times when the airborne network becomes disconnected (e.g., at 1000 UTC). Note that the handover strategy attempts to keep every aircraft at minimum congestion distance from the access network, it does not directly attempt to perfectly balance traffic load among Internet Gateways. Thus, the throughput Θ_{S+H} lies slightly below the theoretical maximum.

8.1.2 Internet gateway A2G capacity vs aircraft set size

Fig. 14 plots the instantaneous ratio of A2G capacity to aircraft set size (C_k/V_k and C_k/N_k) for each Internet Gateway k during the first three hours. With Voronoi cell assignments, some Internet Gateways (e.g., Scotland and Labrador) have plenty of capacity for only a few nodes, whereas others (e.g., Greenland and Iceland) have to serve many aircraft with very little capacity. Thanks to the GLSR handover strategy, each cell breathes aircraft in/out until a congestion equilibrium is reached, overcoming this load/capacity imbalance. In saturation, Internet Gateways serve a number of aircraft roughly proportional to their instantaneous capacity.
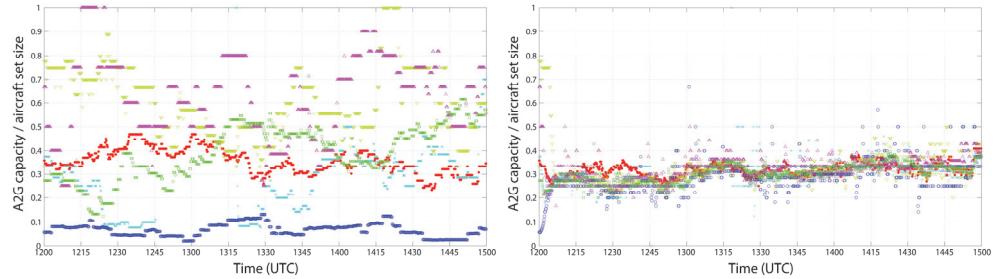
Fig. 14. Instantaneous ratio of A2G capacity to aircraft set size at each Internet Gateway for Voronoi cell assignments (left) and GLSR (right). For each IGW, the color is as in Fig. 15.

Fig. 15 shows the Internet Gateway assignments at 1300 UTC for the G+V and S+H routing schemes. As traffic demand increases, the handover strategy appears to deform the Voronoi diagram by keeping every aircraft at minimum congestion distance from the access network. The trace of traffic through the network is also shown (below), the width of each link indicating the volume of traffic flowing through it. GLSR exploits the rich connectivity of the airborne mesh network, making opportunistic use of the A2G path diversity to avoid buffer congestion as traffic demand fluctuates.
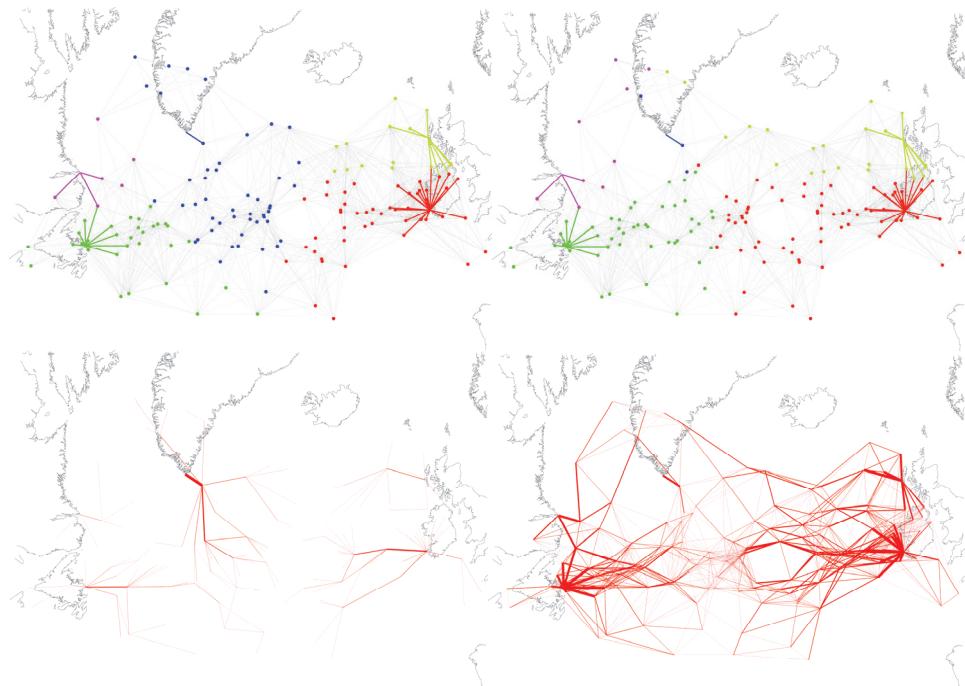

Fig. 15. Internet Gateway assignments and link usage at 1300 UTC for G+V (left) and S+H (right). Width is proportional to link traffic load.

8.2 Results with realistic wireless channel access

In a real aeronautical mesh network, the channel access constraints (c_1)-(c_3) given in Section 3.2 must be satisfied in order to successfully deliver a packet over a radio link. As a result, a link (i,j) will only be able to transmit during a fraction of the frame, as specified in the TDMA schedule, with a capacity $0 \le c_{ij} \le 1$ packets/slot.

8.2.1 Maximum instantaneous throughput

Fig. 16 shows the maximum per-aircraft throughput over the first three hours for the routing schemes defined in Section 7, without interference ($\gamma_o = 0$) and with interference ($\gamma_o = 5$). As a result of interference constraints being taken into account during link scheduling, the variance in A2G capacity among different Internet Gateways is lower. Thus, the distance between the curves μ_{S+V} and μ_{max} is reduced. Regardless of the degree of spatial reuse in the network, the S+H routing scheme approaches the maximum theoretical instantaneous throughput μ_{max} by sharing the total A2G capacity available at any given time among all airborne nodes.
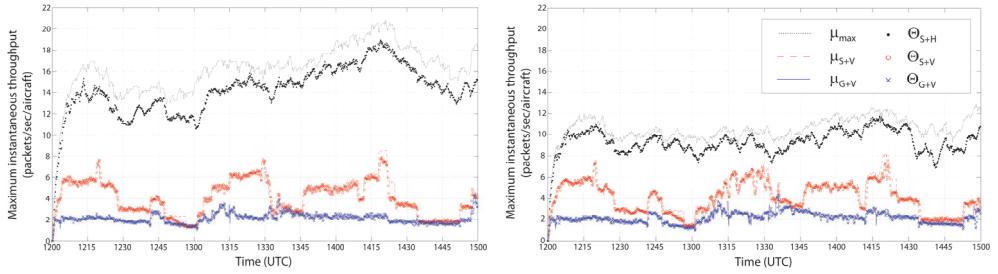

Fig. 16. Maximum instantaneous per-aircraft throughput with γ_o=0 (left) and γ_o=5 (right).

We define the figure of merit χ_R for each routing scheme R as

$$\chi_R = \frac{\int_W \Theta_R(t)\, dt}{\int_W \mu_{max}(t)\, dt} \qquad (30)$$

where the integral is over the simulated time window W, in this case from 1200 UTC to 1500 UTC. Table 2 gives the figures of merit for each routing scheme under the three channel access settings simulated.

	χ_{G+V}	χ_{S+V}	χ_{S+H}
ideal	0.1119	0.2041	0.8930
γ_o=0	0.1063	0.2142	0.8672
γ_o=5	0.2022	0.3876	0.8744

Table 2. Figures of merit for each routing scheme.

Fig. 17 shows the average per-aircraft throughput $\Theta(\lambda)$ and packet delivery ratio $\zeta(\lambda)$ (i.e., the number of packets successfully delivered divided by the number of packets generated) as a function of the per-aircraft traffic demand λ. The two plots are related by

$$\zeta(\lambda) = \frac{\Theta(\lambda)}{\lambda} \tag{31}$$

The curves shown correspond to the routing schemes G+V and S+H under various interference scenarios, and represent the average for 10 static network topologies, equally spaced between 1200 UTC and 1500 UTC (i.e., one topology every 20 minutes).

The interference constraints impact the spatial reuse in the network and therefore the ability to simultaneously schedule A2G links, which pose the traffic bottlenecks in the network. The maximum throughput achievable by the S+H routing scheme is inherently constrained by the total A2G capacity available to the airborne network, which depends on the degree of spatial reuse.
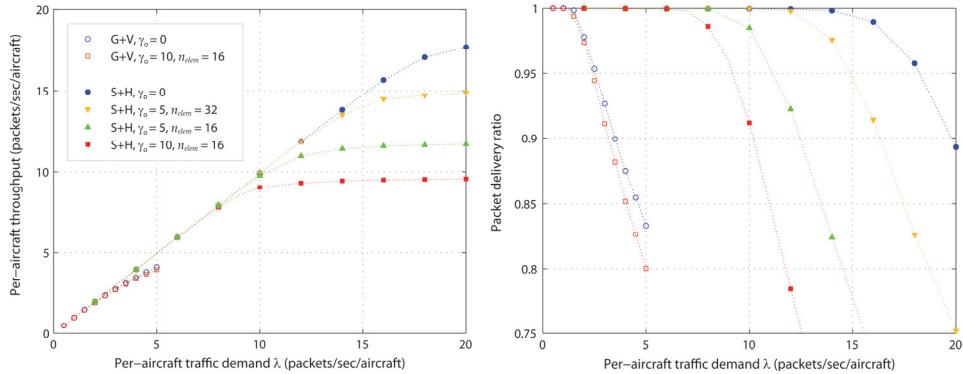

Fig. 17. Per-aircraft throughput and packet delivery ratio as a function of traffic load.

On the other hand, the throughput performance of the G+V routing scheme is relatively insensitive to the reduction in total A2G capacity ensuing from a decrease in spatial reuse, since it does not attempt to exploit the total A2G capacity in the first place.

8.2.2 End-to-end packet delay

Another important performance measure is end-to-end packet delay, defined as the time between the arrival of a packet at the source (Internet Gateway) and its successful reception at the destination (aircraft). Fig. 18 shows the histograms of end-to-end packet delay for $\lambda = 1$ to 10 packets/sec/aircraft under the G+V and S+H routing schemes (with and without interference). These have been obtained for the static network topology at 1200 UTC.

Thanks to the opportunistic nature of GLSR, even at high traffic loads ($\lambda = 10$), almost all packets arrive at their destination aircraft within less than 250 ms (the one-way end-to-end propagation delay for a geostationary satellite link). This is so even though traffic is being routed on a best effort basis, without QoS guarantees.

By contrast, the G+V routing scheme fails to recognize congestion and leads to increased queueing delay and buffer overflow at the bottleneck links, ignoring free available capacity elsewhere in the network. This is responsible for the long tails in the histogram.

Fig. 19 shows the mean of the delay histograms obtained for the G+V and S+H routing schemes as a function of the per-aircraft traffic demand λ under different interference scenarios. As before, the values plotted correspond to the average over 10 static network topologies equally spaced between 1200 UTC and 1500 UTC.
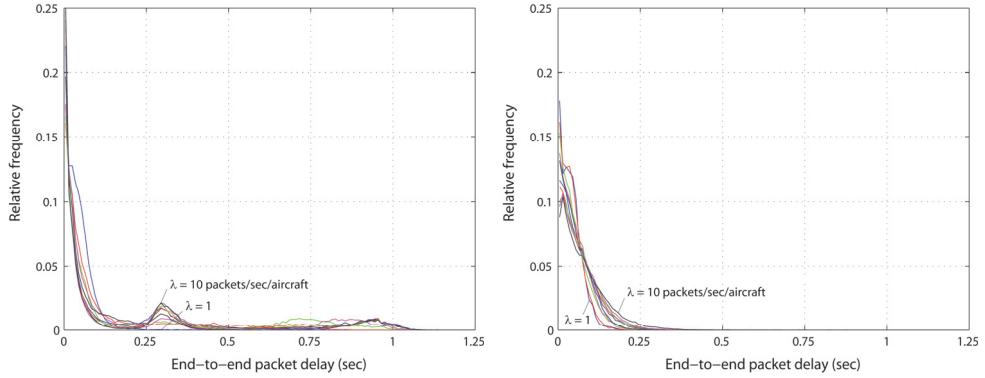
Fig. 18. Delay histograms for G+V (left) and S+H (right) routing schemes at 1200 UTC (γ_o=0).
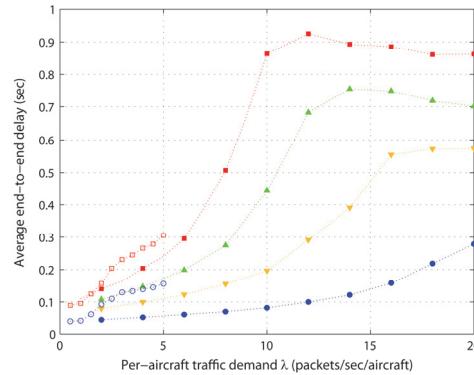

Fig. 19. Average end-to-end packet delay (see legend in Fig. 17).

9. Conclusion

The North Atlantic Corridor constitutes the most interesting scenario for a real deployment of airborne mesh networking technology to provide faster and cheaper inflight internet connectivity during oceanic flight than is currently possible via satellite. In the so-called Airborne Internet, all internet traffic enters/leaves the airborne mesh network via a time-varying number of short-lived air-to-ground (A2G) links, which consequently pose a capacity bottleneck, limiting the maximum data throughput that can be offered to each user (aircraft). Thus, it is essential that the routing strategy keep a balance between the capacity and traffic load of each A2G link. Achieving this balance with minimal overhead in a highly mobile network where link capacity and traffic demand are constantly fluctuating is a challenging task. Our proposed solution, Geographic Load Share Routing (GLSR), requires only the exchange of the aircraft's position, and reacts quickly to fluctuations in traffic demand and link capacity by using instantaneous buffer size information local to the forwarding node. Our simulation results using realistic transatlantic air traffic underscore the importance of a load balancing strategy for the Airborne Internet and confirm GLSR's ability to share the total A2G bandwidth fairly among all airborne users. By exploiting the

full capacity available at each access point and adaptively resizing their geographic jurisdiction to account for congestion, GLSR achieves a per-user throughput close to 90% of the maximum theoretical. This is in stark contrast to the performance of shortest path routing, with a throughput below 20% of the maximum.

10. Acknowledgment

11. References

Ahn, S.; Kim, Y.; Lim, Y. & Lee, J. (2005). Load Balancing in MANET with Multiple Internet Gateways, *IETF Internet Draft*, draft-ahn-manet-multigateway-00, October 2005

Akyildiz, I. & Wang, X. (2005). A survey on wireless mesh networks. *IEEE Communications Magazine*, Vol. 43, No. 9, September 2005, pp. 23-30

Balanis, C. A. (2005). *Antenna Theory: Analysis and Design, 3rd edition*, Wiley-Interscience, April 2005, ISBN 978-0471667827

Bhobe, A. U. & Perini, P. L. (2001). An Overview of Smart Antenna Technology for Wireless Communication, *Proceedings of IEEE Aerospace Conference*, Big Sky, MT, USA, March 2001

Bibb Cain, J.; Billhartz, T.; Foore, L.; Althouse, E. & Schlorff, J. (2003). A Link Scheduling and Ad hoc Networking Approach using Directional Antennas. *Proceedings of IEEE MILCOM 2003*, Vol. 22, No. 1, October 2003

Brännström, R.; Ahlund, C. & Zaslavsky, A. (2005). Maintaining Gateway Connectivity in Multi-hop Ad hoc Networks, *Proceedings of IEEE WLN 2005*, Tampa, FL, USA, November 2005

Chen, D. & Varshney, P. (2007). A Survey of Void Handling Techniques for Geographic Routing in Wireless Networks. *IEEE Communications Surveys and Tutorials*, 2007, pp. 50-67

DirecNet. (2007). DirecNet Task Force Open Session, San Diego, February 2007

Grönkvist, J. (2005). *Interference-based Scheduling in Spatial Reuse TDMA*, Ph.D. Thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2005

He, T.; Stankovic, J. A.; Lu, C. & Abdelzaher, T. (2003). SPEED: A Stateless Protocol for Real-Time Communication in Sensor Networks, *Proceedings of the 23rd IEEE International Conference on Distributed Computing Systems (ICDCS'03)*, USA, May 2003

Hoffmann, F.; Medina, D. & Wolisz, A. (2011). Optimization of Routing and Gateway Allocation in Aeronautical Ad Hoc Networks Using Genetic Algorithms, *Proceedings of IWCMC 2011*, Istanbul, Turkey, July 2011

Huang, C.; Lee, H. & Tseng, Y. (2003). A Two-Tier Heterogeneous Mobile Ad hoc Network Architecture and Its Load Balance Routing Problem, *Proceedings of IEEE VTC 2003 Fall*, Orlando, FL, USA, October 2003

International Air Transport Association. (2007). IATA Schedule Reference Service (SRS). Available from: *http://www.iata.org/ps/publications/srs/*

International Civil Aviation Organization (ICAO). (2008). *North Atlantic Minimum Navigation Performance Specifications (MNPS) Airspace Operations Manual, Edition 2008*, published on behalf of the North Atlantic Systems Planning Group (NAT SPG) by the European and North Atlantic Office of ICAO, August 2008

International Telecommunications Union. (1986). *Recommendation ITU-R P. 528-2, Propagation Curves for Aeronautical Mobile and Radionavigation Services using the VHF, UHF and SHF bands*, ITU, Geneva, Switzerland, 1986

Iordanakis, M.; Yannis, D.; Karras, K.; Bogdos, G.; Dilintas, G.; Amirfeiz, M.; Colangelo, G. & Baiotti, S. (2006). Ad-hoc Routing Protocol for Aeronautical Mobile Ad-Hoc Networks, *Proceedings of 5th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)*, July 2006

Li, L.; Halpern, J. Y.; Bahl, P.; Wang, Y-M. & Wattenhofer, R. (2005). A Cone-Based Distributed Topology-Control Algorithm for Wireless Multi-Hop Networks. *IEEE/ACM Transactions on Networking*, Vol. 13, No. 1, February 2005, pp. 147-159

Mauve, M.; Widmer, J. & Hartenstein, H. (2001). A Survey on Position-based Routing in Mobile Ad Hoc Networks. *IEEE Network*, Vol. 15, No. 6, November 2001, pp. 30-39

McNary, W. (2007). Transformational Aircraft Communication Using a Broadband Mesh Network, *Proceedings of 7th ICNS Conference*, May 2007

Medina, D.; Hoffmann, F.; Ayaz, S. & Rokitansky, C.-H. (2008a). Feasibility of an Aeronautical Mobile Ad Hoc Network Over the North Atlantic Corridor, *Proceedings of IEEE SECON 2008*, San Francisco, CA, USA, June 2008

Medina, D.; Hoffmann, F.; Ayaz, S. & Rokitansky, C.-H. (2008b). Topology Characterization of High Density Airspace Aeronautical Ad Hoc Networks, *Proceedings of IEEE MASS 2008*, Atlanta, GA, USA, September 2008

Medina, D.; Hoffmann, F.; Rossetto, F. & Rokitansky, C.-H. (2010). A Crosslayer Geographic Routing Algorithm for the Airborne Internet, *Proceedings of IEEE ICC 2010*, Cape Town, South Africa, May 2010

Moser, C. (2004). *Ad Hoc Networking with Beamforming Antennas: Modeling, Visualization and Connectivity*, Diploma Thesis, Technical University of Munich (TUM), Munich, Germany, December 2004

Nelson, R. & Kleinrock, L. (1985). Spatial TDMA: A Collision-Free Multihop Channel Access Protocol. *IEEE Transactions on Communications*, Vol. 33, No. 9, September 1985, pp. 934-944

OMNeT++. (2011). Available from: *http://www.omnetpp.org*

Sakhaee, E. & Jamalipour, A. (2006). The Global In-Flight Internet. *IEEE Journal on Selected Areas in Communications*, September 2006

Sakhaee, E. ; Jamalipour, A. & Kato, N. (2006). Aeronautical Ad Hoc Networks, *Proceedings of IEEE WCNC 2006*

Sun, Y.; Belding-Royer, E. M. & Perkins, C. E. (2002). Internet Connectivity for Ad Hoc Mobile Networks. *International Journal of Wireless Information Networks*, Vol. 9, No. 2, April 2002

Tu, H. D. & Shimamoto, S. (2009). Mobile Ad-Hoc Network Based Relaying Data System for Oceanic Flight Routes in Aeronautical Communications. *International Journal of Computer Networks and Communications (IJCNC)*, Vol. 1, No. 1, April 2009

List of Authors

Angeloluca Barba SELEX ELSAG S.p.A.
angeloluca.barba@selexelsag.com

Antonietta Stango Center for TeleInFrastruktur (CTIF), Aalborg, Denmark
as@es.aau.dk

Bertrand Noharet Acreo AB, Sweden
Bertrand.Noharet@acreo.se

Chris Roeloffzen University of Twente, the Netherlands
C.G.H.Roeloffzen@ewi.utwente.nl

Christian Kissling German Aerospace Center (DLR), Institute of Communications
and Navigation, Oberpfaffenhofen, Germany
Christian.Kissling@dlr.de

Daniel Medina German Aerospace Center (DLR), Institute of Communications
and Navigation, Oberpfaffenhofen, Germany
Daniel.Medina@dlr.de

David Marpaung University of Twente, the Netherlands
D.A.I.Marpaung@ewi.utwente.nl

Edward Hall ITT Corporation, Ft. Wayne, Indiana, USA
ward.hall@itt.com

Elias Pschernig University of Salzburg
epschern@cosy.sbg.ac.at

Eriza Hafid Fazli TriaGnoSys GmbH, Germany
eriza.fazli@triagnosys.com

Federica Battisti Università degli Studi Roma TRE
federica.battisti@uniroma3.it

Felix Hoffmann German Aerospace Center (DLR), Institute of Communications
and Navigation, Oberpfaffenhofen, Germany
Felix.Hoffmann@dlr.de

Frederic Durand SITA, Aircom Service Definition, France
fred.durand@sita.aero

Jaco Verpoorte National Aerospace Laboratory (NLR), the Netherlands
Jaco.Verpoorte@nlr.nl

James M. Budinger National Aeronautics and Space Administration (NASA), Glenn
Research Center, Cleveland, Ohio, USA
james.m.budinger@nasa.gov

Joachim Szodruch German Aerospace Center (DLR), Cologne, Germany
Joachim.Szodruch@dlr.de

Kai Xu School of Engineering, Design and Technology, University of
Bradford, Bradford, UK
K.J.Xu@Bradford.ac.uk

Luc Longpre SITA, Aircom IP Product, Canada
luc.longpre@sita.aero

Luigia Micciullo Università di Firenze, Firenze, Italy
luigia.micciullo@unifi.it

Mathieu Gineste	Thales Alenia Space France mathieu.gineste@thalesaleniaspace.com
Matteo Berioli	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Matteo.Berioli@dlr.de
Max Ehammer	University of Salzburg mehammer@cosy.sbg.ac.at
Michael Schnell	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Michael.Schnell@dlr.de
Muhammad Ali	School of Engineering, Design and Technology, University of Bradford, Bradford, UK M.Ali28@Bradford.ac.uk
Muhammad Muhammad	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Muhammad.Muhammad@dlr.de
Neeli Prasad	Center for TeleInFrastruktur (CTIF), Aalborg, Denmark np@kom.aau.dk
Nicolas Van Wambeke	Thales Alenia Space France nicolas.van-wambeke@thalesaleniaspace.com
Nikolaos Fistas	EUROCONTROL nikolaos.fistas@eurocontrol.int
Norbert Rapacz	AGH University of Science and Technology, Kracow, Poland rapacz@kt.agh.edu.pl
Oliver Lücke	TriaGnoSys GmbH, Germany oliver.luecke@triagnosys.com
Piotr Pacyna	AGH University of Science and Technology, Kracow, Poland pacyna@kt.agh.edu.pl
Prashant Pillai	School of Engineering, Design and Technology, University of Bradford, Bradford, UK p.pillai@bradford.ac.uk
Rens Baggen	IMST GmbH, Germany baggen@imst.de
Richard Degenhardt	German Aerospace Center (DLR), Institute of Composite Structures and Adaptive Systems, Brunswick, Germany Richard.Degenhardt@dlr.de
Romano Fantacci	Università di Firenze, Firenze, Italy romano.fantacci@unifi.it
Simon Plass	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Simon.Plass@dlr.de
Snjezana Gligorevic	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Snjezana.Gligorevic@dlr.de
Thomas Gräupl	University of Salzburg tgraeupl@cosy.sbg.ac.at

Tomaso de Cola	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Tomaso.DeCola@dlr.de
Tomasz Chmielecki	AGH University of Science and Technology, Kracow, Poland tomasz.chmielecki@gmail.com
Tommaso Pecorella	Università di Firenze, Firenze, Italy tommaso.pecorella@unifi.it
Ulrich Epple	German Aerospace Center (DLR), Institute of Communications and Navigation, Oberpfaffenhofen, Germany Ulrich.Epple@dlr.de
Willem Beeker	LioniX BV, the Netherlands W.P.Beeker@lionixbv.nl
Yifang Liu	School of Engineering, Design and Technology, University of Bradford, Bradford, UK y.liu18@bradford.ac.uk
Yim Fun Hu	School of Engineering, Design and Technology, University of Bradford, Bradford, UK y.f.hu@bradford.ac.uk
Yongqiang Cheng	School of Engineering, Design and Technology, University of Bradford, Bradford, UK y.cheng4@bradford.ac.uk